

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 20:26:41 ; Search time 18 Seconds  
(without alignments)  
717.411 Million cell updates/sec

Title: US-10-015-385A-194  
Perfect score: 1374  
Sequence: 1 MGLSIFLLCVLGLSQATP.....GVYTYICKYVDWIRMNRN 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1374	100.0	248	1	KLK8_HUMAN
2	630.5	45.9	260	1	KLK8_HUMAN
3	622.5	45.3	260	1	NRPN_MOUSE
4	621.5	45.2	260	1	NRPN_RAT
5	618.5	45.0	250	1	KLK8_HUMAN
6	599	43.6	256	1	KLK8_HUMAN
7	569.5	41.4	248	1	TRY2_CHICK
8	569.5	41.4	251	1	KLK5_HUMAN
9	568.5	41.4	250	1	KLK9_HUMAN
10	567	41.3	277	1	KLK9_HUMAN
11	566	41.2	248	1	TRY2_CHICK
12	565	41.1	248	1	TRY1_CHICK
13	564.5	41.1	293	1	KLK5_HUMAN
14	564	41.0	276	1	KLK8_HUMAN
15	563	41.0	244	1	TRY2_XENLA
16	561	40.8	231	1	TRY1_PIG
17	561	40.8	243	1	TRY1_BOVIN
18	561	40.8	244	1	KLK6_HUMAN
19	555	40.4	243	1	TRY1_XENLA
20	555	40.4	253	1	KLK7_HUMAN
21	553	40.2	246	1	TRY2_MOUSE
22	550	40.0	247	1	TRY2_BOVIN
23	546	39.7	247	1	TRY2_CANFA
24	545	39.7	246	1	TRY1_CANFA
25	544.5	39.6	247	1	TRY3_RAT
26	543	39.5	238	1	TRY1_RAT
27	540	39.3	236	1	TRY3_SALSA
28	533.5	38.8	247	1	TRY4_RAT
29	533	38.8	246	1	TRY2_RAT
30	527.5	38.4	261	1	KLK7_RAT
31	521.5	38.0	246	1	KLK_PIG
32	521.5	38.0	261	1	KLK6_MOUSE
33	519	37.8	246	1	TRYA_RAT

34	518	37.7	246	1	TRYB_RAT
35	518	37.7	261	1	KLK8_RAT
36	513.5	37.4	231	1	TRY2_SALSA
37	513.5	37.4	239	1	KLK2_CAVPO
38	510.5	37.2	261	1	KLK3_MOUSE
39	508	37.0	242	1	TRY1_SALSA
40	505	36.8	260	1	ESTA_CANFA
41	503	36.6	263	1	KLK8_PRANA
42	502	36.5	259	1	KLK8_RAT
43	501	36.5	247	1	TRY2_HUMAN
44	499.5	36.4	241	1	TRYX_GADMO
45	498	36.2	254	1	KLK4_HUMAN

## ALIGNMENTS

RESULT 1  
KLK8\_HUMAN STANDARD; PRT; 248 AA.  
AC Q9UKR0; Q9UKR1;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Kallikrein 12 precursor (EC 3.4.21.-) (Kallikrein-like protein 5)  
DE (KLK-L5).  
GN KLK12 OR KLK15.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=20118156; PubMed=10652563;  
RA Yousef G.M., Luo L.-Y., Diamandis E.P.;  
RT "Identification of novel human kallikrein-like genes on chromosome 19q13.3-q13.4";  
RL Anticancer Res. 19:2843-2852(1999).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RA Yousef G.M., Magklara A., Scorilas A., Diamandis E.P.;  
RT "Cloning of new alternatively spliced forms of the kallikrein-like gene 5 (KLK-L5).";  
RN [3]  
RP Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=20510030; PubMed=11054574;  
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J., Moss P., Paepker B., Wang K.;  
RT "Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";  
RN [5]  
RP Gene 257:119-130(2000).  
RN [6]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V., Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S., Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Barnes J., Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S., Andreise T., Frankel M., Attix C., Amico-Keller G., Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B., Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S., Carraro A.V.;  
RT "Sequence analysis of chromosome 19q13.4";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Secreted (Probable).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q9UKR0-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q9UKR0-2; Sequence=VSP\_005403;  
CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily. P32821

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AF135025; AAD26426.2; -  
DR EMBL; AF135025; AAF06065.1; -  
DR EMBL; AF243527; AAG33365.1; -  
DR EMBL; AC011473; AAG23258.1; -  
DR HSSP; PO0763; LDPO.  
DR MEROPS; S01.020; -  
DR Genew; HGNC:6360; KLK12.  
DR MIM; 605539; -  
DR GO; GO:0005576; C:extracellular; NAS.  
DR GO; GO:0004252; F:serine-type endopeptidase activity; NAS.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00089; trypsin; 1  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; TRYPSIN; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.

KW Hydrolase, Serine protease, Glycoprotein, Signal;  
KW Alternative splicing;  
FT SIGNAL 1 17  
FT CHAIN 18 248  
FT ACT\_SITE 62 62  
FT ACT\_SITE 108 108  
FT ACT\_SITE 200 200  
FT DISULFID 28 161  
FT DISULFID 47 63  
FT DISULFID 133 235  
FT DISULFID 140 206  
FT DISULFID 172 186  
FT DISULFID 196 222  
FT CARBOHYD 24 24  
FT CARBOHYD 163 163  
FT VARSPLIC 236 248

FT SEQUENCE 248 AA; 26733 MW; BB473E98F8BAF703 CRC64;  
Query Match 100.0%; Score 1374; DB 1; Length 248;  
Best Local Similarity 100.0%; Pred No. 1.7e-105;  
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLSIFLLCVLGLSQAATPKIPNGTECGRNSQPQWQVGLFEGTSRCGGVLIDHRWLTA 60  
Db 1 MGLSIFLLCVLGLSQAATPKIPNGTECGRNSQPQWQVGLFEGTSRCGGVLIDHRWLTA 60

QY 61 AHCSGRYVRLGEHSLSDWTEQIRHSGFSVTHPGYLGASTSHEDRLRLRLPVRV 120  
Db 61 AHCSGRYVRLGEHSLSDWTEQIRHSGFSVTHPGYLGASTSHEDRLRLRLPVRV 120

QY 121 TSSVQPLPLNDCATAGTECHVSGWGTTHNRPFPDQLQCLNLSIVSHATCHGVYPGRI 180  
Db 121 TSSVQPLPLNDCATAGTECHVSGWGTTHNRPFPDQLQCLNLSIVSHATCHGVYPGRI 180

QY 181 TSNMVCAGGVPQDACCQGGPLVCGVLQGLVSWGSGVPCQDQIPGYTIVICKYVDW 240  
Db 181 TSNMVCAGGVPQDACCQGGPLVCGVLQGLVSWGSGVPCQDQIPGYTIVICKYVDW 240

QY 241 IRIMMRN 248  
Db 241 IRIMMRN 248

RESULT 2  
ID KLB8 HUMAN STANDARD; PRT; 260 AA.  
AC OG0259; Q9HCB3; Q9UIL9; Q9UQ47;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Neuropilin precursor (BC 3.4.21.-) (NP) (Kallikrein 8) (Ovasin) (Serine  
DE protease TADG-14) (Tumor-associated differentially expressed gene-14  
DE protein).  
GN KLB8 OR PRS19 OR TADG14 OR NRPN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Hippocampus;  
RX MEDLINE=98372070; PubMed=9714609;  
RA Yoshida S., Taniguchi M., Hirata A., Shiosaka S.;  
RT "Sequence analysis and expression of human neuropilin cDNA and gene.";  
RL Gene 213:9-16(1998).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RC TISSUE=Brain;  
RX MEDLINE=93203457; PubMed=10102990;  
RA Mitsui S., Tsuruoka N., Yamashiro K., Nakazato H., Yamaguchi N.;  
RT "A novel form of human neuropilin, a brain-related serine protease, is  
RT generated by alternative splicing and is expressed preferentially in  
RT human adult brain.";  
RL Eur. J. Biochem. 260:627-634(1999).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Ovary;  
RX MEDLINE=99413504; PubMed=10485494;  
RA Underwood L.J., Tanimoto H., Wang Y., Shigemasa K., Farley T.H.,  
RA O'Brien T.J.;  
RT "Cloning of tumor-associated differentially expressed gene-14, a novel  
RT serine protease overexpressed by ovarian carcinoma.";  
RL Cancer Res. 59:4435-4439(1999).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RA Gan L., Gellinas R., Gown A.M., Moss P., Smith R., Wang K.;  
RT "Molecular cloning and characterization of a novel serine protease,  
RT ovasin, a potential molecular marker for ovarian carcinomas.";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=20510030; PubMed=11054574;  
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,  
RA Moss P., Paepker B., Wang K.;  
RT "Sequencing and expression analysis of the serine protease gene  
RT cluster located in chromosome 19q13 region.";  
RL Gene 257:119-130(2000).  
RN [6]  
RP SEQUENCE OF 1-164 FROM N.A. (ISOFORM 1).  
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,  
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,  
RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Garnes J.,  
RA Dangnan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,  
RA Andreise T., Frankheim M., Attix C., Amico-Keller G., Cosfield J.,  
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,  
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,  
RA Olsen A.S., Carrano A.V.;  
RT "Sequence analysis of chromosome 19q13.4";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Suggested to be involved in kindling epileptogenesis and  
CC hippocampal plasticity.  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-, Lys-.

CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;

```

CC CC      IsoId=060259-1; Sequence=Displayed;
CC CC      Name=2;
CC CC      IsoId=060259-2; Sequence=VSP_005401;
CC CC      -!- TISSUE SPECIFICITY: Isoform 1 is predominantly expressed in the
CC CC      pancreas while isoform 2 is expressed in adult brain and
CC CC      hippocampus. Both forms are also found in fetal brain and
CC CC      placenta. Not detected in kidney, spleen, liver and lung.
CC CC      -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
CC CC      -----
CC CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC      the European Bioinformatics Institute. There are no restrictions on its
CC CC      use by non-profit institutions as long as its content is in no way
CC CC      modified and this statement is not removed. Usage by and for commercial
CC CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC      or send an email to license@isb-sib.ch).
CC CC      -----
CC CC      EMBL; AB009849; BAA28673.1; -
CC CC      EMBL; AB012761; BAA28676.1; -
CC CC      EMBL; AB010780; BAA88684.1; -
CC CC      EMBL; AB008390; BAA82665.1; -
CC CC      EMBL; AB008927; BAA82666.1; -
CC CC      EMBL; AF055982; AAD56050.1; -
CC CC      EMBL; AF095742; AAD25979.1; -
CC CC      EMBL; AF095743; AAD29574.1; -
CC CC      EMBL; AF243527; AAG33361.1; -
CC CC      EMBL; AC011473; AAG23254.1; -
CC CC      HSSP; O61955; INPM.
CC CC      MEROPS; S01.244; -
CC CC      Genew; HGNC:6369; KUK8.
CC CC      MIM; 605644; -
CC CC      GO; GO:0008236; F:serine-type peptidase activity; TAS.
CC CC      GO; GO:0007399; P:neurogenesis; TAS.
CC CC      InterPro; IPR009003; Cys Ser trypsin.
CC CC      InterPro; IPR001254; Peptidase_S1.
CC CC      InterPro; IPR001314; Peptidase_S1A.
CC CC      Pfam; PF00089; trypsin; 1.
CC CC      PRINTS; PR00722; CHYMOTRYPSIN.
CC CC      SMART; SM00020; Tryp_SPC; 1.
CC CC      PROSITE; PS00240; TRYPSIN_DOM; 1.
CC CC      PROSITE; PS00134; TRYPSIN_HIS; 1.
CC CC      PROSITE; PS00135; TRYPSIN_SER; 1.
CC CC      Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
CC CC      Alternative splicing.
CC CC      SIGNAL 1 28 POTENTIAL.
CC CC      PROPEP 29 32 BY SIMILARITY.
CC CC      CHAIN 33 260 NEURPSIN.
CC CC      ACT_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CC      ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CC      ACT_SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CC      DISULFID 39 173 BY SIMILARITY.
CC CC      DISULFID 58 74 BY SIMILARITY.
CC CC      DISULFID 145 246 BY SIMILARITY.
CC CC      DISULFID 152 218 BY SIMILARITY.
CC CC      DISULFID 184 198 BY SIMILARITY.
CC CC      DISULFID 208 233 BY SIMILARITY.
CC CC      CARBOHYD 110 110 N-LINKED (GLCNAC...) (POTENTIAL).
CC CC      VARSPIC 23 23 A -> AAGSLDILTKVAENLPVHLNPMQPSQSHCPRG
CC CC      WRSNLPAPRA (in isoform 2).
CC CC      /FTID=VSPF_005401.
CC CC      SEQUENCE 260 AA; 28048 MW; EF43955B8C83B660 CRC64;
CC CC      Query Match 45.9%; Score 630.5; DB 1; Length 260;
CC CC      Best Local Similarity 50.2%; Pred. No. 1.1e-44;
CC CC      Matches 123; Conservative 24; Mismatches 93; Indels 5; Gaps 3;
CC CC      5 IFLLL---CVILGSAATPKINFTGCRNSOPWQGLPEGTSRLCGVLDHRWVLTAA 61
CC CC      13 MFLLLGGAWAGHSRAQEDKVLGHECPHSPQQAALFQGCQLLCGGVILVGNVLTAA 72
CC CC      62 HCSGSRVYVRLGHEISQLDWTQETRHGSFVTHPGYLGAST-SHEHDLRLRLPLPVY 120
CC CC      73 HCKPKYTVRLGDHSLQNKDGPEQEPVQVQSPHPCYNSSDVEDNHDMLQLRQDASL 132

```

```

QY 121 TSSVQPLPLPNDCATAGTECHVSGKGTNHRNPPDPLLOCLNLSIVSHATCHGVYPGRI 180
DB 133 GSKVKPISLADHCTQPGQKQTVSGMGTVTSPRENFPDTLNCVAEVKIFPKKCEDAYPQOI 192
QY 181 TSNMVCAGVPGQACQDSDGGPLVCGGVLOGLVSWGSGVPGCGDGIQGVYTYICKYVDW 240
DB 193 TDGMVCAGSSKGATCCQDSDGGPLVCDALQGITWSGS-DFGGRSDKPGVYTNICRYLDW 251
QY 241 IRMIM 245
DB 252 IKXII 256

RESULT 3
NRPN MOUSE
ID NRPN MOUSE STANDARD; PRT; 260 AA.
AC Q61955;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neutropin precursor (EC 3.4.21.-) (NP (Kallikrein 8)).
GN KUK8 OR PRSS19 OR NRPN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Hippocampus;
RX MEDLINE=953348817; PubMed=7623137;
RA Chen Z.-L., Yoshida S., Kato K., Momota Y., Suzuki J., Tanaka T.,
RA Ito J., Nishino H., Aimoto S., Kiyama H., Shiosaka S.;
RT "Expression and activity-dependent changes of a novel limbic-serine
RT protease gene in the hippocampus.";
RL J. Neurosci. 15:5088-5097(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Yoshida S., Hirata A., Inoue N., Shiosaka S.;
RT "Cloning and assignment of mouse neutropin gene, Prss19 to chromosome
RT 7B4.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Czech II; TISSUE=Mammary gland;
RX MEDLINE=22380257; PubMed=12477932;
RA Klausner R.B., Feingold E.A., Grouse L.H., Derge J.G.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Matra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF N-TERMINUS, AND CHARACTERIZATION.
RC STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=98225202; PubMed=9556608;
RA Shimizu C., Yoshida S., Shibata M., Kato K., Momota Y., Matsumoto K.,
RA Shiosaka T., Midorikawa R., Kamachi T., Kawabe A., Shiosaka S.;
RT "Characterization of recombinant and brain neutropin, a

```





RP SEQUENCE FROM N.A.  
RC STRAIN=Fischer; TISSUE=Brain;  
RX MEDLINE=98389725; PubMed=9722524;  
RA Davies B.J., Pickard B.S., Steel M., Morris R.G.M., Lathe R.;  
RT "Serine proteases in rodent hippocampus";  
RL J. Biol. Chem. 273:23004-23011(1998)  
CC -!- FUNCTION: Suggested to be involved in kindling epileptogenesis and  
CC hippocampal plasticity. Has a strong proteolytic activity against  
CC fibronectin (by similarity).  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-, Lys-.  
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).  
CC -!- TISSUE SPECIFICITY: Restricted to hippocampus.  
CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.  
CC  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AJ005641; CAA06643.1; -.  
CC HSP; O61955; 1NPM.  
CC MEROPS; S01.244; -.  
CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC SMART; SM00020; TRYPSIN; 1.  
CC PROSITE; PS02040; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00135; TRYPSIN\_SSR; 1.  
CC Hydrolase; Serine protease; Glycoprotein; Zymogen; signal.  
KW SIGNAL 1 28 POTENTIAL.  
FT PROPEP 29 32 BY SIMILARITY.  
FT CHAIN 33 260 NEUROPSIN.  
FT ACT\_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 39 173 BY SIMILARITY.  
FT DISULFID 58 74 BY SIMILARITY.  
FT DISULFID 145 246 BY SIMILARITY.  
FT DISULFID 152 218 BY SIMILARITY.  
FT DISULFID 184 198 BY SIMILARITY.  
FT DISULFID 208 233 BY SIMILARITY.  
FT CARBOHYD 110 110 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 260 AA; 28510 MW; 58DF4F0602A0B7F5 CRC64;  
Query Match 45.2%; Score 621.5; DB 1; Length 260;  
Best Local Similarity 48.8%; Pred. No. 6.1e-44;  
Matches 118; Conservative 32; Mismatches 90; Indels 3; Gaps 3;  
QY 5 IFLL-CVGLGSAATPKIFNGTEGRNSQFQWGLFEGTSLRCGGVLDHRVWLTAAHC 63  
DB 15 LFLMGAWAGLTRAQSKILEGQCKPHSQPWQTALFQGERLVCVGLVGDVWLTAAHC 74  
QY 64 SGRYVRLGHELSQLDWTQTRHSGFVTHPGYLGAS-TSHEHLRLRLRLPVRVTS 122  
DB 75 KDKYSVRLGDHSLQKRDPEQRIQVARSSTQHPFCNFPNSDESHDMLRLQNSANLGD 134  
QY 123 SVQPLPLNDCAATGTECHVSGWGIHTHPNPDPDLLOCLNLISVSHATGCHVYGRITS 182  
DB 135 KVPIELANLCPKVQKCIISGWTGVTSPQENFENTLNCAEVKIYSQNKERAYPGKITE 194  
QY 183 NMVCAAGVPGQDCAQDGGGLVCGVGLQGLVSWGSGVPGQDGPVGYTCYKYYDWIR 242  
DB 195 GWCAGSSNGADTCQDSSGGLVCGVGLQGLITWGS-DPCQKPKGVYTKICRYTNWIK 253  
QY 243 MTM 245

Db 254 KTM 256  
RESULT 5  
KLKB HUMAN STANDARD; PRT; 250 AA.  
AC Q9UBX7; O75837; Q9NS65;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Kallikrein 11 precursor (EC 3.4.21.-) (Hippostasin) (Trypsin-like  
DE protease).  
GN KLK11 OR PRSS20 OR TLSP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Hippocampus;  
RX MEDLINE=98438738; PubMed=9765601;  
RA Yoshida S., Ianiguchi M., Suemoto T., Oka T., He X.P., Shiosaka S.;  
RT "cDNA cloning and expression of a novel serine protease, TLSP";  
RL Biochim. Biophys. Acta 1399:225-228(1998).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RC TISSUE=Hippocampus, and Prostate;  
RX MEDLINE=20329229; PubMed=10872828;  
RA Mitsu S., Yamada T., Okui A., Kominami K., Uemura H., Yamaguchi N.;  
RT "A novel isoform of a kallikrein-like protease, TLSP/hippocastin,  
RT (PRSS20), is expressed in the human brain and prostate.";  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=20130117; PubMed=10662548;  
RA Yousef G.M., Scorrilas A., Diamantis E.P.;  
RT "Genomic organization, mapping, tissue expression, and hormonal  
RT regulation of trypsin-like serine protease (TLSP PRSS20), a new  
RT member of the human kallikrein gene family.";  
RL Genomics 63:88-96(2000).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=20510030; PubMed=11054574;  
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,  
RA Moss P., Paepier B., Wang K.;  
RT "Sequencing and expression analysis of the serine protease gene  
RT cluster located in chromosome 19q13 region.";  
RL Gene 257:119-130(2000).  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,  
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,  
RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Barnes J.,  
RA Dangnan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,  
RA Andreise T., Frankheim M., Attix C., Amico-Keller G., Coefield J.,  
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,  
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,  
RA Olsen A.S., Carrano A.V.;  
RT "Sequence analysis of chromosome 19q13.4";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Testis;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Vallah D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smalls M.A.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: Possible multifunctional protease. Efficiently cleaves  
 CC bz-Phe-Arg-4-methylcoumarinyl-7-amide, a kallikrein substrate, and  
 CC weakly cleaves other substrates for kallikrein and trypsin.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q9UBX7-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q9UBX7-2; Sequence=VSP\_005402;  
 CC -!- TISSUE SPECIFICITY: Expressed in brain, skin and prostate. Isoform  
 CC 1 is expressed preferentially in brain; isoform 2 in prostate.  
 CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB012917; BAA33404.1; AUT\_INIT.  
 DR EMBL; AB013730; BAA88713.1; -.  
 DR EMBL; AB041036; BAA96797.1; -.  
 DR EMBL; AF146623; AAD47815.1; -.  
 DR EMBL; AF243527; AAG33364.1; -.  
 DR EMBL; AC011473; AAG33257.1; -.  
 DR EMBL; BC022068; AAH22068.1; -.  
 DR HSSP; P00763; 1DPO.  
 DR MEROPS; S01\_257; -.  
 DR Genew; HGNC:6359; KLK11.  
 DR MiW; 60434; -.  
 DR GO; GO:0008236; F:serine-type peptidase activity; TAS.  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR01254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS02040; TRYPsin\_DOM; 1.  
 DR PROSITE; PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE; PS00135; TRYPsin\_SER; 1.  
 DR Hydrolase; Serine protease; Glycoprotein; Signal; Zymogen;  
 KW Alternative splicing  
 FT SIGNAL 1 18  
 FT PROPEP 19 21  
 FT CHAIN 22 250  
 FT ACT\_SITE 62 62  
 FT ACT\_SITE 110 110  
 FT ACT\_SITE 203 203  
 FT DISULFID 28 163  
 FT DISULFID 47 63  
 FT DISULFID 135 237  
 FT DISULFID 142 209  
 FT DISULFID 174 188  
 FT DISULFID 199 224  
 FT CARBOHYD 99 99  
 FT CARBOHYD 165 165  
 FT CARBOHYD 181 181  
 FT CARBOHYD 210 210  
 FT VARSPIC 1 1  
 M -> MQLRLRLDMKSSGRGLTAKEPGARSSPLQAM

FT (in isoform 2).  
 FT /FTId=VSP\_005402.  
 SQ SEQUENCE 250 AA; 27466 MW; 192D910BBCDC7A56 CRC64;  
 Query Match 45.0%; Score 618.5; DB 1; Length 250;  
 Best Local Similarity 48.2%; Pred. No. 1e-43;  
 Matches 120; Conservative 35; Mismatches 89; Indels 5; Gaps 4;  
 QY 3 LSTFLLLCVLGSLQAATPKIFNGTEGRNSQPWQVGLFEGTSRLCGVLDHRLVLTAAH 62  
 DB 4 LQILLLALATGLVGGGT-RIIKGFECKPHSPQWQAALFEKTRLLCGATLIAPRWLLTAAH 62  
 QY 63 CSSRSYVRLRGHLSQLDQTEQIRHSFGSVTHPGYLGA--STSHEHDLRLRLRLPVRV 120  
 DB 63 CLKPRYIVHLGQNLNQBEGCEQTRATSFPHFGFNLSLPNKDHRNDMLVQWASPVSI 122  
 QY 121 TSSVQPLPLPNDCAATAGTCHVSGWGITNHRNPFDPDLLOCLNLSIVSHATGCVTPGRI 180  
 DB 123 TWAVRPLTLSSRCVTAGTSLISGWSGTSPPQLRPLHTLRCAITITIEHQKCNAYPGNI 182  
 QY 181 TSNMVCAG-CVPGQDACQDGGGLVCGVLOGLVSGSVGPGCGDGIPIGVYTIKCYVD 239  
 DB 183 TDTMVCASVQEGKQSGDGGPLVNCOSLQGIISWGQ-DPCAITRKPGVYTKVKCYVD 241  
 QY 240 WIRMENNN 248  
 DB 242 WIQETMKN 250  
 RESULT 6  
 KLKF HUMAN STANDARD; PRT; 256 AA.  
 ID \_KLKF\_HUMAN Q9H2R3; Q9H2R4; Q9H2R6; Q9HBG9;  
 AC Q9H2R5; O15358; Q9H2R3; Q9H2R4; Q9H2R6; Q9HBG9;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, last sequence update)  
 DT 28-FEB-2003 (Rel. 41, last annotation update)  
 DE Kallikrein 15 precursor (BC 3.4.21.-) (ACO protease).  
 GN KKL15.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.  
 RX PubMed=11010966;  
 RA Yousef G.M., Scorilas A., Jung K., Ashworth L.K., Diamandis E.P.;  
 RT "Molecular cloning of the human kallikrein 15 gene (KLK15). Up-  
 RT regulation in prostate cancer.";  
 RL J. Biol. Chem. 276:53-61(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20510030; PubMed=11054574;  
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,  
 RA Moss P., Paepker B., Wang X.;  
 RT "Sequencing and expression analysis of the serine protease gene  
 RT cluster located in chromosome 19q13 region.";  
 RL Gene 257:119-130(2000).  
 RN [3]  
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=94289486; PubMed=8018728;  
 RA Dihanich M.E., Spiess M.;  
 RT "A novel serine proteinase-like sequence from human brain.";  
 RL Biochim. Biophys. Acta 1218:225-228(1994).  
 CC -!- FUNCTION: Protease whose physiological substrate is not yet known.  
 CC -!- SUBCELLULAR LOCATION: Secreted (Probable).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=4;  
 CC Name=1;  
 CC IsoId=Q9H2R5-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q9H2R5-2; Sequence=VSP\_005405;  
 CC Name=3;

```
CC      IsoId=Q9H2R5-3; Sequence=VSP_005406, VSP_005407;
CC      Name=4;
CC      IsoId=Q9H2R5-4; Sequence=VSP_005404;
CC      TISSUE SPECIFICITY: Highest expression in the thyroid gland. Also
CC      expressed in the prostate, salivary, and adrenal glands and in the
CC      colon testis and kidney.
CC      -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AF242195; AAG09469.1; -
CC      EMBL; AF242195; AAG09470.1; -
CC      EMBL; AF242195; AAG09471.1; -
CC      EMBL; AF242195; AAG09472.1; -
CC      EMBL; AF243527; AAG33354.1; -
CC      EMBL; X75363; CAA53145.1; ALT_SEQ.
CC      HSP; P00763; IDPO.
CC      MEROPS; S01.081; -.
CC      GO; GO:000576; C:extracellular; NAS.
CC      GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
CC      GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
CC      InterPro; IPR009003; Cys_Ser trypsin.
CC      InterPro; IPR001254; Peptidase S1.
CC      InterPro; IPR001314; Peptidase_S1A.
CC      Pfam; PF00089; trypsin; 1.
CC      PRINTS; PR00722; CHYMOTRYPSIN.
CC      PROSITE; PS50240; TRYPSIN_DOM; 1.
CC      PROSITE; PS00134; TRYPSIN_HIS; 1.
CC      PROSITE; PS00135; TRYPSIN_SER; FALSE NEG.
CC      Hydrolase; Serine protease; Glycoprotein; Signal; Zymogen;
KW      Alternative splicing.
FT      SIGNAL      1   16   POTENTIAL.
FT      PROPEP     17   21   ACTIVATION PEPTIDE (POTENTIAL).
FT      CHAIN      22   256  KALLIKREIN 15.
FT      ACT_SITE   62   62   CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      ACT_SITE  106   106  CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      ACT_SITE  209   209  CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      CARBOHYD   171   171  N-LINKED (GLCNAC. .) (POTENTIAL).
FT      CARBOHYD   232   232  N-LINKED (GLCNAC. .) (POTENTIAL).
FT      VARSPLIC   122   206  Missing (in isoform 4).
FT      VARSPLIC   122   256  Missing (in isoform 2).
FT      VARSPLIC   161   161  V -> G (in isoform 3).
FT      VARSPLIC   162   256  /FTid=VSP_005406.
FT      VARSPLIC   147   160  Missing (in isoform 3).
FT      CONFLICT   147   160  SHNEPGTAGSPRSQ -> PLSSP (IN REF. 2).
FT      SEQUENCE   256 AA; 28087 MW; B5EBF8D6022786B5 CRC64;
Query Match      43.6%; Score 599; DB 1; Length 256;
Best Local Similarity 47.1%; Pred. No. 4.2e-42;
Matches 121; Conservative 32; Mismatches 88; Indels 16; Gaps 6;
QY      5  IFLLCVLGLSQAT--PKIFNGTCGRNSQPWQVGLFEGTSLRCGGVLIDHRVWLTAAH 62
DB      3  LLTLTSLASTAAQDQGLLEGDECAFPQWQVALYERCFNCGASLLSPHWLSAAH 62
QY      63  CSGRYVVRIGELHSLSQDWTQIRHSGFVTHPGYLGASTSHEDHRLRLRLPVRVTS 122
DB      63  CQSRFMRVRLGELHNLRLKRDGPEQLTRTSRVIPHPRY--EARSHRNDIMLLRLVQPARLNP 120
QY      123  SVQPLPLPNDCATAGTCHVSGMGIYNH-----PRN--PPDLLQCLNLSVSHATC 172
DB      121  QVRPAVLPTRCPPHGEACVSGWGLVSHNPEGTAGSPRSQVSLPDTLLHCANISIIISTSC 180
QY      173  HGVPYGRITSNMVCAGGV-PQDACQGDGGGLVCGGLVQGLVSWGSGVPCGQDGPVGY 231
```

```
DB      181  DKSPGRLTNMVCAGRGARSGCEGSGGLVCGGLVQGLVSWGVDV-PCDNTTKPGVY 239
QY      232  TYICKYVDWIRMIMRNN 248
DB      240  TKVCHYLEWIRETMKRN 256
RESULT 7
TRY3_CHICK
ID      TRY3_CHICK  STANDARD;      PRT;      248 AA.
AC      Q90629;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Trypsin II-P29 precursor (EC 3.4.21.4).
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC      Gallus.
OX      NCBI_TaxID=9031;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Pancreas;
RX      MEDLINE=95251611; PubMed=7733885;
RA      Wang K., Gan L., Lee I., Hood L.E.;
RT      Isolation and characterization of the chicken trypsinogen gene
      family. ";
RL      Biochem. J. 307:471-479(1995).
CC      -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC      -!- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
CC      -!- SUBCELLULAR LOCATION: Extracellular.
CC      -!- TISSUE SPECIFICITY: High levels are seen in the pancreas while
CC      lower levels are found in the liver, spleen and thymus.
CC      -!- SIMILARITY: Belongs to peptidase family S1.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; U15157; AAA79914.1; -
CC      FTR; S55066; S55066.
CC      HSP; P00763; IDPO.
CC      MEROPS; S01.151; -.
CC      InterPro; IPR009003; Cys_Ser trypsin.
CC      InterPro; IPR001254; Peptidase S1.
CC      InterPro; IPR001314; Peptidase_S1A.
CC      Pfam; PF00089; trypsin; 1.
CC      PRINTS; PR00722; CHYMOTRYPSIN.
CC      SMART; SM00020; Tryp_SPC; 1.
CC      PROSITE; PS50240; TRYPSIN_DOM; 1.
CC      PROSITE; PS00134; TRYPSIN_HIS; 1.
CC      PROSITE; PS00135; TRYPSIN_SER; 1.
CC      Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;
KW      Calcium-binding; Signal; Multigene family.
FT      SIGNAL      1   16   BY SIMILARITY.
FT      PROPEP     17   25   ACTIVATION PEPTIDE (BY SIMILARITY).
FT      CHAIN      26   248  TRYPSIN II-P29.
FT      ACT_SITE   65   65   CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      METAL      77   77   CALCIUM (BY SIMILARITY).
FT      METAL      79   79   CALCIUM (VIA CARBONYL OXYGEN)
      (BY SIMILARITY).
FT      METAL      82   82   CALCIUM (VIA CARBONYL OXYGEN)
      (BY SIMILARITY).
FT      METAL      87   87   CALCIUM (BY SIMILARITY).
FT      ACT_SITE  109   109  CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      ACT_SITE  202   202  CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      DISULFID   32   162  BY SIMILARITY.
FT      DISULFID   50   66   BY SIMILARITY.
```

FT DISULFID 134 235 BY SIMILARITY.  
FT DISULFID 141 208 BY SIMILARITY.  
FT DISULFID 173 187 BY SIMILARITY.  
FT DISULFID 198 222 BY SIMILARITY.  
FT SITE 196 196 REQUIRED FOR SPECIFICITY (BY SIMILARITY).  
SQ SEQUENCE 248 AA; 26622 MW; E5E16B07622B588E CRC64;  
  
Query Match 41.4%; Score 569.5; DB 1; Length 248;  
Best Local Similarity 47.5%; Pred. No. 1e-39;  
Matches 116; Conservative 33; Mismatches 78; Indels 11; Gaps 6;  
  
QY 5 IFLLLCVLGLS-----QAATPKIFNGTECGRNQPMQVGLFEGTSURCGGVLDHRLVLR 59  
DB 4 LFLILSLGAAVAPPGADDDKIVGGTCPEHSVYQVLSNGVHF-CGGLINSQWVLS 62  
  
QY 60 AAHCGSRYWVRIGEHLSQLDWTEQIRHGSFVTHPGYLGASTSHEDHLRLRLPVR 119  
DB 63 AAHCYKSRIOVRIGEXNIDVQEDSEVRSSVIRHPKY--SSITLNDMLKLASAVE 120  
  
QY 120 VTSVQPLPNDPCATAGTECHVSGWGTINHPNPFDDLLQLCLMLSTIVSHATCHGVYVGR 179  
DB 121 YSADIQIPALPSSCAKAGTECLISGNGYLSNGYNYPELLOCLNAPILSDQEQEAYFGD 180  
  
QY 180 ITSNMVCAGGVP--QDACCQSGSGPLVCGGVGLQGLVSGVSGCGQDGPVGYTYICKYV 238  
DB 181 ITSNMVCVGFLEGGKDSQSGSGPVGNGELQGVSWG-IG-CALXGYPGVYTKVNCYV 238  
  
QY 239 DWIR 242  
DB 239 DWIQ 242  
  
RESULT 8  
ID KLKE\_HUMAN STANDARD; PRT; 251 AA.  
AC Q9P0G3;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Kallikrein 14 precursor (EC 3.4.21.-) (Kallikrein-like protein 6)  
DE (KLK-L6).  
GN KLK14 OR KLK6.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yousef G.M., Diamantis E.P.;  
RT "Molecular characterization, mapping, and tissue expression of KLK6,  
a hormonally regulated kallikrein-like gene."  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
RX MEDLINE=21250997; PubMed=1152573;  
RA Hooper J.D., Bui L.T., Rae F.K., Harvey T.J., Myers S.A.,  
RA Ashworth L.K., Clements J.A.;  
RT "Identification and characterization of KLK14, a novel kallikrein  
serine protease gene located on human chromosome 19q13.4 and expressed  
in prostate and skeletal muscle."  
RL Genomics 73:117-122(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,  
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,  
RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Barnes J.,  
RA Danganan L., Eriker A., Christensen M., Georgescu A., Avila J., Liu S.,  
RA Andreise T., Trankheim M., Attix C., Amico-Keller G., Coefield J.,  
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,  
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,  
RA Olsen A.S., Carrano A.V.;  
RT "Sequence analysis of chromosome 19q13.4."  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

RN [4]  
RP TISSUE SPECIFICITY.  
RX MEDLINE=20545474; PubMed=10969073;  
RA Harvey T.J., Hooper J.D., Myers S.A., Stephenson S.A., Ashworth L.K.,  
RA Clements J.A.;  
RT "Tissue-specific expression patterns and fine mapping of the human  
kallikrein (KLK) locus on proximal 19q13.4."  
RL J. Biol. Chem. 275:37397-37406(2000).  
CC -!- SUBCELLULAR LOCATION: Secreted (Probable).  
CC -!- TISSUE SPECIFICITY: High expression in brain, bone marrow and  
fetal liver. Also expressed in liver, pancreas, fetal spleen,  
prostate and skeletal muscle.  
CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-----  
CC EMBL; AF161221; AAD50773.2; -;  
DR EMBL; AF283669; AAK48523.1; -;  
DR EMBL; AF283670; AAK48524.1; -;  
DR EMBL; AC011473; AAC23260.1; -;  
DR HSP; P00763; IDPO.  
DR MEROPS; S01.029; -;  
DR Genew; HGNC:6362; KLK14.  
DR MIM; 606135; -;  
DR GO; GO:0005576; C:extracellular; NAS.  
DR GO; GO:0004252; P:serine-type endopeptidase activity; NAS.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_SPc; 1.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Signal; Zymogen.  
FT SIGNAL 1 18 POTENTIAL.  
FT PROPEP 19 24 ACTIVATION PEPTIDE (POTENTIAL).  
FT CHAIN 25 251 KALLIKREIN 14.  
FT ACT\_SITE 67 67 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 111 111 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 204 204 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 31 164 BY SIMILARITY.  
FT DISULFID 52 68 BY SIMILARITY.  
FT DISULFID 143 210 BY SIMILARITY.  
FT DISULFID 175 189 BY SIMILARITY.  
FT DISULFID 200 225 BY SIMILARITY.  
SQ SEQUENCE 251 AA; 27452 MW; 9087953BAPAFED25 CRC64;  
  
Query Match 41.4%; Score 569.5; DB 1; Length 251;  
Best Local Similarity 46.9%; Pred. No. 1e-39;  
Matches 119; Conservative 33; Mismatches 87; Indels 15; Gaps 6;  
  
QY 5 IFLLLCVLGLS-----SQATPKIFNGTECGRNQPMQVGLFEGTSUR--CGGVLDHRL 55  
DB 1 MFLLLTALQVLAIAMTQSQDENKIIGGTCRSPQWQAALLPGRFRCGGLSGQ 60  
  
QY 56 VWLTAHSGSRYWVRIGEHLSQLDWTEQIRHGSFVTHPGYLGASTSHEDHLRLRLR 115  
DB 61 WVTTAHCRPTLQVALGKHLRWEATQVLRVVRQVTHPNY--NSRTHDNDMLLQLQ 118  
  
QY 116 LPVUTSSVQPLPNDPCATAGTECHVSGWGTINHPNPFDDLLQLCLMLSTIVSHATCHGV 175  
DB 119 QPARIGRAVRPEVTQACASPGTSCVSGWGTISSPIARYPASLQCVNINISPDVCQKA 178  
  
QY 176 YPGRITSNMVCAGGVP--QDACCQSGSGPLVCGGVGLQGLVSGVSGCGQDGPVGYTY 233

Db 179 YPRITTEGMVCA-GVPGGKDSQGGSLVCRQLQGLVSMG-MERCALPGYGVYTN 236  
Qy 234 ICKYVDWIRMIRN 247  
Db 237 LCKYRSWIETMRD 250

RESULT 9  
KLK9 HUMAN  
ID\_KLK9\_HUMAN STANDARD; PRT; 250 AA.  
AC Q9UKR3; Q9Y433;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Kallikrein 9 precursor (EC 3.4.21.-) (Kallikrein-like protein 3) (KLK-L3).  
DE L3).  
GN KLK9.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20118156; PubMed=10652563;  
RA Yousef G.M., Iuo L.-Y., Diamandis E.P.;  
RT "Identification of novel human kallikrein-like genes on chromosome 19q13.3-q13.4.";  
RL Anticancer Res. 19:2843-2852(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20247258; PubMed=10783266;  
RA Yousef G.M., Diamandis E.P.;  
RT "The expanded human kallikrein gene family: locus characterization and molecular cloning of a new member, KLK-L3.";  
RL Genomics 65:184-194(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20510030; PubMed=11054574;  
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J., Moss P., Paepel B., Wang K.;  
RT "Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";  
RL Gene 257:119-130(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V., Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S., Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Garnes J., Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S., Andreise T., Trankheim M., Attix C., Amico-Keller G., Cosfield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmüller B., Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S., Carrano A.V.;  
RT "Sequence analysis of chromosome 19q13.4.";  
RL Submitted (OCT-2000) to the EMBL/Genbank/DDay databases.  
CC -!- SUBCELLULAR LOCATION: Secreted (Probable).  
CC -!- TISSUE SPECIFICITY: Skin, thymus, trachea, cerebellum and spinal cord.  
CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF135026; AAD26427.2; -.  
DR EMBL; AF243527; AAG33362.1; -.  
DR EMBL; AC011473; AAG32355.1; -.  
DR HSSP; P00763; 1DPO.

DR MEROPS; S01.307; -.  
DR Genew; HGNC:6370; KLK9.  
DR MIM; 605504; -.  
DR GO; GO:0005576; C:extracellular; NAS.  
DR GO; GO:0004252; F:serine-type endopeptidase activity; NAS.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.  
DR InterPro; IPR009003; CysSer trypsin.  
DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolyase; Serine protease; Glycoprotein; Signal.  
FT SIGNAL 1 15  
FT CHAIN 16 250 KALLIKREIN 9.  
FT ACT\_SITE 63 63 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 111 111 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 204 204 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 29 164 BY SIMILARITY.  
FT DISULFID 48 64 BY SIMILARITY.  
FT DISULFID 136 238 BY SIMILARITY.  
FT DISULFID 143 210 BY SIMILARITY.  
FT DISULFID 175 189 BY SIMILARITY.  
FT DISULFID 200 225 BY SIMILARITY.  
FT CARBOHYD 131 131 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 211 211 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 250 AA; 27512 MW; F2785245B063E98B CRC64;

Query Match 41.4%; Score 568.5; DB 1; Length 250;  
Best Local Similarity 46.7%; Pred. No. 1.3e-39;  
Matches 119; Conservative 32; Mismatches 89; Indels 15; Gaps 7;

Qy 1 MGLSIFILLCVLGLSQAATPKIFNGTECGRNSQPQVQLFEGTSLRCGVLIDHRWLTA 60  
Db 3 LGLICALLSLLAGHWADT-RAIGAECECFNSQPQAGLFLHTRFCGATLSDRWLTA 61  
Qy 61 AHCSGRYRWVRLGEHSLQSDWTQIRHSGSVT----HPGY---LGASTSHEHDLRLR 113  
Db 62 AHCRKPYLWVRLGEHHLWKWEGPEQL-----PRVTDFFPHGFKDLSAN-DHNDIMLR 116  
Qy 114 LRLPVRTSSVQPLPLPNDCATAGTECHVSGWGITNHPNPFDDLLQCLNLSIVSHATCH 173  
Db 117 LPROARLSPAVQPLNLSQTCVSPGMCCLISGAVSPKALFPVTLQCANISLENKLCH 176  
Qy 174 GVPYGRITSNMVCAG-GVPGQDACCQDGGSLVCRQLQGLVSMGSGVPCGQDGIPGYVT 232  
Db 177 WAYPGHISDMLCAGLWEGGRSGCGDGGSLVCRQLQGLVSMGSGVPCGQDGIPGYVT 235  
Qy 233 YICKYVDWIRMIRN 247  
Db 236 SVCHYLDWIQIEMEN 250

RESULT 10  
KLK9 HUMAN  
ID\_KLK9\_HUMAN STANDARD; PRT; 277 AA.  
AC Q9UKR3; Q9Y433;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Kallikrein 13 precursor (EC 3.4.21.-) (Kallikrein-like protein 4) (KLK-L4).  
DE L4).  
GN KLK13 OR KLK14.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=20229789; PubMed=10766816;  
RA Yousef G.M., Chang A., Diamandis E.P.;  
RT "Identification and characterization of KLK-L4, a new kallikrein-like  
RL gene that appears to be down-regulated in breast cancer tissues.";  
RN J. Biol. Chem. 275:11891-11898(2000).  
[2]  
RP SEQUENCE FROM N.A.  
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,  
RA Burkhardt-Schulz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,  
RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Garnes J.,  
RA Danganan L., Eriar A., Christensen M., Georgescu A., Avila J., Liu S.,  
RA Andreise T., Frankheim M., Attix C., Amico-Keller G., Coefield J.,  
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,  
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,  
RA Olsen A.S., Carrano A.V.;  
RT "Sequence analysis of chromosome 19q13.4";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
[3]  
RN SEQUENCE OF 1-180 FROM N.A.  
RC TISSUE=Uterus;  
RA Ansong W., Winkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Secreted (Probable).  
CC -!- TISSUE SPECIFICITY: Expressed in prostate, breast, testis and  
CC salivary gland.  
CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF135024; AAD26425.2; -;  
CC EMBL; AC011473; AAG23259.1; -;  
CC EMBL; AL050220; CAB43320.1; ALT\_INIT.  
CC HSP; P00763; IDPO.  
CC MEROPS; S01.306; -;  
CC Genew; HGNC:6361; KLK13.  
CC MIM; 605505; -;  
CC GO; GO:0005576; C:extracellular; NAS.  
CC GO; GO:0004252; F:serine-type endopeptidase activity; NAS.  
CC GO; GO:0006508; P:proteolysis and peptidolysis; NAS.  
CC InterPro; IPR009003; Cys Ser trypsin.  
CC InterPro; IPR001254; Peptidase S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC SMART; SM00020; Tryp SP; 1.  
CC PROSITE; PS50240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00135; TRYPSIN\_SRR; 1.  
KW Hydrolase; Serine protease; Glycoprotein; Signal.  
FT SIGNAL 1 16  
FT CHAIN 17 277 KALLIKREIN 13.  
FT ACT\_SITE 76 76 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 124 124 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 218 218 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 42 178 BY SIMILARITY.  
FT DISULFID 61 77 BY SIMILARITY.  
FT DISULFID 157 224 BY SIMILARITY.  
FT DISULFID 189 203 BY SIMILARITY.  
FT DISULFID 214 239 BY SIMILARITY.  
FT CARBOHYD 30 30 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CARBOHYD 225 225 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CONFLICT 170 180 VNPFTKLCAN -> GMHPRWPEAP (IN REF. 3).  
FT SEQUENCE 277 AA; 30570 MW; BA8A9BDCFB5D542 CRC64;  
  

Query Match 41.3%; Score 567; DB 1; Length 277;  
Best Local Similarity 46.8%; Pred. No. 1.9e-39;  
Matches 116; Conservative 36; Mismatches 80; Indels 16; Gaps 6;

QY 13 GLSQRAATPKIFN-----GTEGRNSQWQVGLFEGTSIRCGVLIIDHRWVLTAAH 62  
DB 18 GVSQESS-KVLNTNGTSGFLPGGYTCFPHSQWQALVQGRLLCGVLIHVPKWVLTAAH 76  
QY 63 CGSRVWVRIGHSLSOLDWTEQIRHSQFSTVTHPGVIGAST--SHEHDLRLRLRLRLEVRV 120  
DB 77 CLKEGKLVYGRHALGRVAGQVREVVHSIHPPEYRSPFTHLNHDDIMLLEQSPVQL 136  
QY 121 TSSVQELPLP-NDCAVAGTECHVSGWGI TNHPRNPDPDLLOCLNLSIVSHATCHGVYGR 179  
DB 137 TGYIQTLP LSHNNRLPGTTCRVSGWGTTSPOVNYPKTLQCANIQLRSDDECRQVYPGK 196  
QY 180 ITSNWVCAGVVP-GQACQDSDGGLVCGVLOGVLSWGSVGPCGOGGIPGVVYTIKVV 238  
DB 197 ITDNMLCAGTKEGGKDCSCBDSGGPLVNCNRTLYGVSWGDF-PCGQPDPRGVTVRSRYV 255  
QY 239 DWIRMIIR 246  
DB 256 LWIREIR 263  
  

RESULT 11  
TRY2\_CHICK  
ID\_TRY2\_CHICK STANDARD; PRT; 248 AA.  
AC Q90628;1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, last sequence update)  
DT 01-NOV-1997 (Rel. 42, last annotation update)  
DE Trypsin 1-P38 precursor (EC 3.4.21.4).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pancreas;  
RX MEDLINE=95251611; PubMed=7733885;  
RT Wang K., Gan L., Lee I., Hood L.E.;  
RT "Isolation and characterization of the chicken trypsinogen gene  
RT family.";  
RL Biochem. J. 307:471-479(1995).  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.  
CC -!- COFACTOR: Binds 1 calcium ion per subunit (By similarity).  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- TISSUE SPECIFICITY: High levels are seen in the pancreas while  
CC lower levels are found in the liver, spleen and thymus.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U15156; AAA79913.1; -;  
CC PIR; S55067; S55067.  
CC HSP; P00763; IDPO.  
CC MEROPS; S01.258; -;  
CC InterPro; IPR009003; Cys Ser trypsin.  
CC InterPro; IPR001254; Peptidase S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC SMART; SM00020; Tryp SP; 1.  
CC PROSITE; PS50240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00135; TRYPSIN\_SRR; 1.  
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;  
KW Calcium-binding; Signal; Multigene family.



DE Kallikrein 5 precursor (BC 3.4.21.-) (Stratum corneum tryptic enzyme)  
GN KLK5 OR SCHE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Stratum corneum;  
RX MEDLINE=9445563; PubMed=10514489;  
RA Brattsand M., Egelrud T.;  
RT "Purification, molecular cloning, and expression of a human stratum  
RT corneum trypsin-like serine protease with possible function in  
RT desquamation.";  
RL J. Biol. Chem. 274:30033-30040(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20118156; PubMed=10652563;  
RA Yousef G.M., Luo L.-Y., Diamandis E.P.;  
RT "Identification of novel human kallikrein-like genes on chromosome  
RT 19q13.3-q13.4.";  
RL Anticancer Res. 19:2843-2852(1999).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20510030; PubMed=11054574;  
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,  
RA Moss P., Paepker B., Wang K.;  
RT "Sequencing and expression analysis of the serine protease gene  
RT cluster located in chromosome 19q13 region.";  
RL Gene 257:119-130(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RX MEDLINE=22477932; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Heide F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Greenwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: May be involved in desquamation.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Expressed in skin, breast, brain and testis.  
CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL; AF168768; AAF03101.1; -;  
CC DR EMBL; AF135028; AAD26429.1; -;  
CC DR EMBL; AF243527; AAG33358.1; -;  
CC DR EMBL; BC008036; AAH08036.1; -;  
CC HSSP; P00763; 1DPO.

DR MEROPS; SOL.017; -;  
DR Genew; HGNC:6366; KLK5.  
DR MIM; 605643; -;  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR GO; GO:0004252; F:serine-type endopeptidase activity; NAS.  
DR GO; GO:0008544; P:epidermal differentiation; TAS.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; TRYP\_SPC; 1.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
DR Hydrolase; Serine protease; Glycoprotein; Signal.  
FT SIGNAL 1 22 POTENTIAL.  
FT CHAIN 23 293 KALLIKREIN 5.  
FT ACT\_SITE 108 108 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 153 153 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 245 245 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 73 206 BY SIMILARITY.  
FT DISULFID 93 109 BY SIMILARITY.  
FT DISULFID 178 279 BY SIMILARITY.  
FT DISULFID 185 251 BY SIMILARITY.  
FT DISULFID 217 231 BY SIMILARITY.  
FT DISULFID 241 266 BY SIMILARITY.  
FT CARBOHYD 69 69 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 173 173 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 252 252 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 25 56 MISSING (IN REF. 3).  
SQ SEQUENCE 293 AA; 32020 MW; D92C92F5609E5946 CRC64;  
Query Match 41.1%; Score 564.5; DB 1; Length 293;  
Best Local Similarity 45.3%; Pred. No. 3.2e-39;  
Matches 107; Conservative 41; Mismatches 83; Indels 5; Gaps 4;  
QY 15 SQATPKIFNGTSGRNSQWVG-LFEGTSLRCGGVLDHRVLTAAHCSGRYVRLG 73  
Db 60 SDDSSRIINGSDCMHTQWQAALLRPNQLYCGAVLHPQMLTAAHCRKVKFVRLG 119  
QY 74 EHSLSOLDWTEQIRHSGF-SVTHPGYLGASTSHEHDLRLRLPVRVTSVQPLPLPD 132  
Db 120 HYSLSPVYESGQMFQGVKSIHPGY--SHPGSHNDMLIKLNRIRPTKDVRLVSSH 177  
QY 133 CATAGTECHVSGMGI TNHPRNPPFDLLQCLNLSIVSHATGCHGVYGRITSNMVCAGVPG 192  
Db 178 CPSAGTKCLVSGWGTGKSPQVHPFKVLQCLNISVLISQKRCEDAYPRQIDTDFCAGDKAG 237  
QY 193 QDACQSGGFLYCGGVGLGVSGVPCGQDGPVGVYICKYVDWIRMVNN 248  
Db 238 RDSQSGSGFVVCNGSLQGLVSWGDY-PCARNRPNRVYINLCFTKWIQETQAN 292  
RESULT 14  
KLKA HUMAN  
ID KLKA\_HUMAN STANDARD; PRT; 276 AA.  
AC O43240; Q99920; Q9GZW9;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Kallikrein 10 precursor (BC 3.4.21.-) (Protease serine-like 1) (Normal  
DE epithelial cell-specific 1).  
GN KLK10 OR PRRS1 OR NES1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Epithelium;



RX MEDLINE=96320486; PubMed=8764136;  
 RA Liu X.-L., Wazer D.E., Watanabe K., Band V.;  
 RT "Identification of a novel serine protease-like gene, the expression  
 of which is down-regulated during breast cancer progression.";  
 RL Cancer Res. 56:3371-3379(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99321170; PubMed=9647736;  
 RA Luo L.-Y., Herbrick J.A., Scherer S.W., Beatty B., Squire J.,  
 RA Diamandis E.P.;  
 RT "Structural characterization and mapping of the normal epithelial  
 cell-specific 1 gene.";  
 RL Biochem. Biophys. Res. Commun. 247:580-586(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20510030; PubMed=11054574;  
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,  
 RA Moss P., Paepker B., Wang K.;  
 RT "Sequencing and expression analysis of the serine protease gene  
 cluster located in chromosome 19q13 region.";  
 RL Gene 257:119-130(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Lanerding J.E., McCready P.M., Skowronski E., Viewanathan V.,  
 RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,  
 RA Phan H., Velasco N., Do L., Regala M., Terry A., Brower A., Ganes J.,  
 RA Danganan L., Eriar A., Christensen M., Georgescu A., Avila J., Liu S.,  
 RA Andreise T., Trankheim M., Attix C., Amico-Keller G., Coefield J.,  
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,  
 RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,  
 RA Olsen A.S., Carrano A.V.;  
 RT "Sequence analysis of chromosome 19q13.4";  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RT TISSUE=Uterus;  
 RX MEDLINE=22389257; PubMed=12477932;  
 RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.N., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [6]  
 RP CHARACTERIZATION.  
 RX MEDLINE=99025848; PubMed=9809976;  
 RA Goyal J., Smith K.M., Cowan J.M., Wazer D.E., Lee S.W., Band V.;  
 RT "The role for NES1 serine protease as a novel tumor suppressor";  
 RL Cancer Res. 58:4782-4786(1998).  
 CC -!- FUNCTION: Has a tumor-suppressor role for NES1 in breast and  
 CC prostate cancer.  
 CC -!- SUBCELLULAR LOCATION: Secreted (Probable).  
 CC -!- TISSUE SPECIFICITY: Expressed in breast, ovary and prostate.  
 CC -!- DEVELOPMENTAL STAGE: Down-regulated during breast cancer  
 CC progression.  
 CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF024605; AAB81602.1; -  
 DR EMBL; AF055481; AAC14266.1; -  
 DR EMBL; AF243527; AAG33363.1; -  
 DR EMBL; AC011473; AAG23256.1; -  
 DR EMBL; BC002710; AAH02710.1; -  
 DR HSP; P00763; 1DPO.  
 DR MEROPS; S01.246; -  
 DR Genew; HGNC:6358; KKL10.  
 DR MIM; 602673; -  
 DR GO; GO:0005576; C:extracellular; TAS.  
 DR GO; GO:0008236; F:serine-type peptidase activity; TAS.  
 DR InterPro; IPR009003; Cys Ser trypsin.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00240; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; FALSE NEG. Signal; Anti-oncogene.  
 KW Hydrolase; Serine protease; Glycoprotein; Signal; Anti-oncogene.  
 FT SIGNAL 1 30  
 FT CHAIN 31 276  
 FT KALLIKREIN 10.  
 FT ACT\_SITE 86 86  
 FT ACT\_SITE 137 137  
 FT ACT\_SITE 229 229  
 FT DISULFID 52 162  
 FT DISULFID 71 87  
 FT DISULFID ? 263  
 FT DISULFID 169 235  
 FT DISULFID 201 215  
 FT DISULFID 225 250  
 FT CARBOHYD 39 39  
 FT CONFLICT 50 50  
 FT CONFLICT 149 149  
 FT CONFLICT P -> L (IN REF. 3 AND 4).  
 SQ SEQUENCE 276 AA; 30138 MW; 82A2507379BAB313 CRC64;  
 Query Match 41.0%; Score 564; DB 1; Length 276;  
 Best Local Similarity 45.9%; Pred No. 3.3e-39;  
 Matches 118; Conservative 29; Mismatches 92; Indels 18; Gaps 5;  
 QY 7 LLLCVLGLSQA-----TPKIFNGTECRNSQPWQVGLFEGTSLRCGGVLIDHRWV 57  
 DB 23 LLMAQLWAAEAALLPQNDRLDPEAY-GAPCARGSQPWQVSLFNGLSFHCAGVLVDQSWV 81  
 QY 58 LTAACSGSRVWVLGEHSLQDLWTEQIRHSGSVTHPGVYGAS-----TSHEDLRL 111  
 DB 82 LTAACGCKPLWARVGDHLLQLLQ-GEQLRRTTSRVHPKYGQSGGPIPRTRDHDHML 140  
 QY 112 LRLRLPVRVTSVQPLPLPNDCATAGTECHVSGWGITNHPNPPFDLLQCLNLSTVSHAT 171  
 DB 141 LKLARPVVPGRVRAQLPYRCAQPGDCCQVAGGTTAARVKYKNGLTCSITILSPKE 200  
 QY 172 CHGVYFGRITSNMVCAGVGPDACQGSQSGPLVCGVQLQGLVSGVSGVPGQGDGPVY 231  
 DB 201 CEVFPYGVVTTNNMICAGLDRCQDPCQSDSGGFLVCDTLOGLISWG-VYPCGSAQHPAVY 259  
 QY 232 TYICKYVDWIRWIMNN 248  
 DB 260 TOICKYMSWINKVRSN 276  
 RESULT 15  
 TRY2\_XENLA  
 ID TRY2\_XENLA STANDARD; PRT; 244 AA.  
 AC P70059;  
 DT 01-NOV-1997 (Rel. 35, Created)

01-NOV-1997 (Rel. 35, Last sequence update)  
10-OCT-2003 (Rel. 42, Last annotation update)  
Trypsin precursor (EC 3.4.21.4)  
Xenopus laevis (African clawed frog).  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
NCBI\_TaxID=8355;  
[1]  
SEQUENCE FROM N.A.  
Wang K., Lytle L., Gan L., Hood L.E.;  
Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
-!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa.  
-!- COFACTOR: Binds 1 calcium ion per subunit (By similarity).  
-!- SUBCELLULAR LOCATION: Extracellular.  
-!- SIMILARITY: Belongs to peptidase family S1.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
-----  
EMBL; U72330; AAB17274.1; -.  
HSP; P00763; IDPO.  
DR DR MEROPS; S01.258; -.  
DR DR InterPro; IPR009003; Cys Ser trypsin.  
DR DR InterPro; IPR001254; Peptidase S1.  
DR DR InterPro; IPR001314; Peptidase\_S1A.  
DR DR Pfam; PF00089; trypsin; 1.  
DR DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR DR SMART; SM00020; TRYPSIN\_SPC; 1.  
DR DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Digestion; Zymogen; Calcium-binding;  
KW Signal; Multigene family.  
FT SIGNAL 1 15 BY SIMILARITY.  
FT PROPEP 16 21 ACTIVATION PEPTIDE (BY SIMILARITY).  
FT CHAIN 22 244 TRYPSIN.  
FT ACT\_SITE 61 61 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT METAL 73 73 CALCIUM (BY SIMILARITY).  
FT METAL 75 75 CALCIUM (VIA CARBONYL OXYGEN)  
FT METAL 83 83 (BY SIMILARITY).  
FT METAL 83 83 CALCIUM (BY SIMILARITY).  
FT ACT\_SITE 105 105 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 198 198 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 28 158 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 46 62 BY SIMILARITY.  
FT DISULFID 130 231 BY SIMILARITY.  
FT DISULFID 137 204 BY SIMILARITY.  
FT DISULFID 169 183 BY SIMILARITY.  
FT DISULFID 194 218 BY SIMILARITY.  
FT SITE 192 192 REQUIRED FOR SPECIFICITY (BY SIMILARITY).  
FT SEQUENCE 244 AA; 26079 MW; C63F29CB3300B323 CRC64;  
Query Match 41.0%; Score 563; DB 1; Length 244;  
Best Local Similarity 45.0%; Pred. No. 3.5e-39;  
Matches 112; Conservative 46; Mismatches 83; Indels 8; Gaps 5;  
  
QY 3 LSIFLLCVLGLSQA--ATPKIFNGTEGRNSQPMQVGLFEGTSIRCGGVLIDHRWLTA 60  
DB 1 MKFLVILVLLGAFAVEDDDKIVGGFTCAKNAPYQVSLNAGYHF--CGGSLNSQWVISA 59  
  
QY 61 AHCSGSRVWVRIGBSLSQLDWTQTQIRHSFSTVTHPGYLGASTSHEHDLRLRLRLFRV 120  
DB 60 AHCKRSRIQVELGHNIALNEGTFQIDSQVKIKHPNY--NSRNLNDIMLIKLTSTTARL 117  
  
QY 121 TSSVQPLPLPNDCTAGTCHVSGNGITNHPNRPFPDLQCLINLSIVSHATCHGVPGRI 180  
DB 118 SANIQSVPLPSACASAGTNCILSGWNTLSSGTNYPDLQLCLINAPILITDSQCSNYPGEI 177

QY 181 TSNMYCAGGVP--GDACCGDSCGGLVCGGVLOGLVSWGSVGPCGQDGPVGYTYICKYVD 239  
DB 178 TKNMFCAGFLAGGKDSGCGSPVVCNGQLQGVVSWGY--GCAQRNYPGVYTKVCNFVT 235  
QY 240 WIRMIMRNN 248  
DB 236 WIQSTISSN 244

Search completed: June 16, 2004, 20:33:11  
Job time : 18 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 01:43:09 ; Search time 4624 seconds  
(without alignments)  
10226.481 Million cell updates/sec

Title: US-10-015-385A-193

Perfect score: 1091

Sequence: 1 caagcagtcaccccttgg.....aggggcaaaaaaaaaa 1091

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rod.\*

36: em.htg.man.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1091	100.0	1091	6	AX697125	Sequence
2	1091	100.0	1091	9	AY358524	Homo sapi
3	510.8	46.8	811	9	BSM808731	Sequence
4	506.4	46.4	840	6	AR337568	Human ser
5	506.4	46.4	840	6	BD130047	Sequence
6	493.2	45.2	619	6	AR263961	Homo sapi
7	373.2	34.2	9120	9	AF164623	Homo sapi
8	373.2	34.2	11820	9	AF135025	Homo sapi
9	373.2	34.2	132323	9	AC011473	Homo sapi
10	373.2	34.2	230000	9	AF243527	Homo sapi
11	368.4	33.8	174724	2	AC140096	Pan trogl
12	368.4	33.8	200792	2	AC130782	Pan trogl
13	361.2	33.1	142334	2	AC073185	Homo sapi
14	297	27.2	405	6	AX070938	Sequence
15	294	26.9	176647	2	AC130188	Papio anu
16	219.4	20.1	1146	6	BD193094	207 human
17	217.8	20.0	1166	6	AR152173	Sequence
18	217.6	19.9	1106	6	AX016287	Sequence
19	217.6	19.9	1106	6	BD137019	Human CAS
20	217.6	19.9	1158	6	AX016289	Sequence
21	217.6	19.9	1158	6	BD137020	Human CAS
22	217.6	19.9	1181	9	AB013730	Homo sapi
23	217.6	19.9	1186	9	AB012917	Homo sapi
24	217.6	19.9	1192	6	AR152174	Sequence
25	217.6	19.9	1204	6	AX358932	Sequence
26	217.6	19.9	1204	6	AX362425	Sequence
27	217.6	19.9	1204	6	AX454622	Sequence
28	217.6	19.9	1204	6	AX464372	Sequence
29	217.6	19.9	1204	6	AX491100	Sequence
30	217.6	19.9	1204	6	AX697101	Sequence
31	217.6	19.9	1204	9	AY359014	Homo sapi
32	217.6	19.9	1213	9	BC022068	Homo sapi
33	217.6	19.9	1301	6	BD091587	Novel ser
34	217.6	19.9	1301	9	AB041036	Homo sapi
35	217.6	19.9	1314	6	AR098430	Sequence
36	217.6	19.9	1314	6	BD130920	Serine pr
37	216	19.8	1191	6	BD139483	Extended
38	210	19.2	1335	6	BD232165	A novel m
39	206.8	19.0	833	6	AR060847	Sequence
40	206.8	19.0	833	6	BD082136	Novel pro
41	200.4	18.4	1213	10	AB016226	Mus muscu
42	200.4	18.4	1256	10	AB016227	Mus muscu
43	200.4	18.4	1323	6	BD091588	Novel ser
44	199.4	18.3	1052	6	AR219287	Sequence
45	198	18.1	1283	9	AF283670	Homo sapi

# ALIGNMENTS

RESULT 1  
AX697125  
LOCUS AX697125  
DEFINITION Sequence 193 from Patent WO0078961.  
ACCESSION AX697125  
VERSION AX697125.1 GI:29498087  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Ferrara, N., Stewart, T.A., Williams, P.M., Baker, K.P., Desnoyers, L.,  
Eaton, D.L., Gao, W.Q., Pan, J., Botstein, D., Fong, S., Goddard, A.,  
Godowski, P.J., Gurney, A.L., Smith, V., Tumas, D., Wood, W.I.,



```
Query Match      100.0%; Score 1091; DB 9; Length 1091;
Best Local Similarity 100.0%; Pred. No. 8.1e-238; Indels 0; Gaps 0;
Matches 1091; Conservative 0; Mismatches 0;

QY 1 CAAGCAGGTCTATCCCTTGGTGTGACCTTCAAGAGAGAGAGAGAGAGAGTGGGGGC 60
DB 1 CAAGCAGGTCTATCCCTTGGTGTGACCTTCAAGAGAGAGAGAGAGAGAGTGGGGGC 60

QY 61 ACAGGAAAGGTGACCTCTGAGATCCCTTTCCCGCAGACTTTGGAAGTACCCACC 120
DB 61 ACAGGAAAGGTGACCTCTGAGATCCCTTTCCCGCAGACTTTGGAAGTACCCACC 120

QY 121 ATGGGGCTCAGCATCTTTTGTCTGTGTCTTGGCTCAGCCAGGACACACCG 180
DB 121 ATGGGGCTCAGCATCTTTTGTCTGTGTCTTGGCTCAGCCAGGACACACCG 180

QY 181 AAGATTTTCAATGGCACTGAGTGTGGGCGTAACTACAGCCGTGGCAGGTGGGTGTTT 240
DB 181 AAGATTTTCAATGGCACTGAGTGTGGGCGTAACTACAGCCGTGGCAGGTGGGTGTTT 240

QY 241 GAGGCAACAGCTCGCTGCGGTGGGGGTGTCCTATTGACCAAGTGGTCTCACAGG 300
DB 241 GAGGCAACAGCTCGCTGCGGTGGGGGTGTCCTATTGACCAAGTGGTCTCACAGG 300

QY 301 GCTCACTGACGGGACGAGTACTGGGTGCGCTGGGGGAACACAGCCTCAGCCAGCTC 360
DB 301 GCTCACTGACGGGACGAGTACTGGGTGCGCTGGGGGAACACAGCCTCAGCCAGCTC 360

QY 361 GACTGACCGAGAGATCGGCACAGCGGTCTCTGTGACCCATCCCGGTACTCTCCAG 420
DB 361 GACTGACCGAGAGATCGGCACAGCGGTCTCTGTGACCCATCCCGGTACTCTCCAG 420

QY 421 GCTCGACGAGCAGCAGCAGCAGCTCGGCTGCTGCGCTGCTGCGCTGCTGCGCTG 480
DB 421 GCTCGACGAGCAGCAGCAGCAGCTCGGCTGCTGCGCTGCTGCGCTGCTGCGCTG 480

QY 481 ACCAGCAGGCTTCAACCCCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 540
DB 481 ACCAGCAGGCTTCAACCCCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 540

QY 541 CAGCTCTCAGGTGGGGATCAGCAACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAG 600
DB 541 CAGCTCTCAGGTGGGGATCAGCAACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAG 600

QY 601 TGCTCAACTCTCCATCTGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 601 TGCTCAACTCTCCATCTGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

QY 661 ACAGCAACATGCTGTGTCAGCGCGCTGCCGGGACAGATGCTGTCAGCGGTGATTTCT 720
DB 661 ACAGCAACATGCTGTGTCAGCGCGCTGCCGGGACAGATGCTGTCAGCGGTGATTTCT 720

QY 721 GGGGGCCCCCTGTGTGTGGGGAGTCTTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 721 GGGGGCCCCCTGTGTGTGGGGAGTCTTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 780

QY 781 CCCTGTGGAACAAGTGGCATCCCTGGAGTCTACACCTATATTTGCAAGTATGGAAGTGG 840
DB 781 CCCTGTGGAACAAGTGGCATCCCTGGAGTCTACACCTATATTTGCAAGTATGGAAGTGG 840

QY 841 ATCCGATGATCATGAGGAACAATGACCTGTTTCTTCCACCTCCACCCCAACCCCTTAA 900
DB 841 ATCCGATGATCATGAGGAACAATGACCTGTTTCTTCCACCTCCACCCCAACCCCTTAA 900

QY 901 CTTGGGTACCCCTCTGGCCCTCAGAGCAACAATATCTCTCCATCATCTTCCCTAGTCTC 960
DB 901 CTTGGGTACCCCTCTGGCCCTCAGAGCAACAATATCTCTCCATCATCTTCCCTAGTCTC 960

QY 961 ACTCTTGTGGCTGGGAATCTTCTGGAATTTAACTCTGCTGCGAGCCCTTCTAAGACCCA 1020
DB 961 ACTCTTGTGGCTGGGAATCTTCTGGAATTTAACTCTGCTGCGAGCCCTTCTAAGACCCA 1020

1021 CGAGCGGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATAAATGAAGAGGGGCAAA 1080
```

```
DB 1021 CGAGCGGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATAAATGAAGAGGGGCAAA 1080
QY 1081 AAAAAAAAAAAAA 1091
DB 1081 AAAAAAAAAAAAA 1091

RESULT 3
HSM08731
LOCUS Homo sapiens mRNA; cDNA DKFP686H1078 (from clone DKFP686H1078).
DEFINITION Homo sapiens mRNA; cDNA DKFP686H1078 (from clone DKFP686H1078).
ACCESSION BX648580
VERSION BX648580.1 GI:34367742
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 811)
AUTHORS Bloeker, H., Boecher, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
CONSTRM The German Human cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2003) MIPs, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFP686H1078) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.

FEATURES
Location/Qualifiers
1..811
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFP686H1078"
/tissue_type="human endometrium carcinoma cell line"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
polyA_signal 702..707
polyA_site 723
ORIGIN
Query Match 46.8%; Score 510.8; DB 9; Length 811;
Best Local Similarity 99.6%; Pred. No. 1.1e-105;
Matches 512; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 578 ACCATATCCCGATCTGCTCCAGTGCCTCAACTCTCCATGCTCCCATGCGACCTGCC 637
DB 226 ACCATATCCCGATCTGCTCCAGTGCCTCAACTCTCCATGCTCCCATGCGACCTGCC 285

QY 638 ATGGTGTGTATCCCGGAGAAATCAGAGCAACATGGTGTGCGCGCGCGTCCCGGGGC 697
DB 286 ATGGTGTGTATCCCGGAGAAATCAGAGCAACATGGTGTGCGCGCGCGTCCCGGGGC 345

QY 698 AGGATGCCCTGCCAGGGTGATTTCTGGGGGCCCTCTGGTGTGGGGAGTCTTCAAGGTC 757
DB 346 AGGATGCCCTGCCAGGGTGATTTCTGGGGGCCCTCTGGTGTGGGGAGTCTTCAAGGTC 405

QY 758 TGGTGTCTGGGGTCTGTGGGGCCCTGTGGACAAGATGGCATCTCTGGAGTCAACCT 817
DB 406 TGGTGTCTGGGGTCTGTGGGGCCCTGTGGACAAGATGGCATCTCTGGAGTCAACCT 465

QY 818 ATATTTGCAAGTATGTGGATCCGATGATCATGAGGAACAACACTGACCTGTTCTCT 877
DB 466 ATATTTGCAAGTATGTGGATCCGATGATCATGAGGAACAACACTGACCTGTTCTCT 525
```

QY 878 CCACCTCCACCCACCCCTTAACCTGGGTACCCCTCTGGCCCTCAGAGCACCAATATCT 937  
Db 526 CCACCTCCACCCACCCCTTAACCTGGGTACCCCTCTGGCCCTCAGAGCACCAATATCT 585  
QY 938 CCTCATCACTCCCTAGCTAGCTCCTCTGTGGCTGGGAACCTCTTGGAACTTTAACT 997  
Db 586 CCTCATCACTCCCTAGCTAGCTCCTCTGTGGCTGGGAACCTCTTGGAACTTTAACT 645  
QY 998 CTTGCCAGCCCTTCTAAGACCCACGAGGGGGTGGAGAACTGTGCAATAGCTCTGGAATA 1057  
Db 646 CTTGCCAGCCCTTCTAAGACCCACGAGGGGGTGGAGAACTGTGCAATAGCTCTGGAATA 705  
QY 1058 AATATAATGAGGAGGGGCAAAAAA 1091  
Db 706 AATATAATGAGGAGGGGCAAAAAA 739  
RESULT 4  
AR337568  
LOCUS  
DEFINITION Sequence 5 from patent US 6566498.  
ACCESSION AR337568  
VERSION AR337568.1 GI:33723963  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 840)  
AUTHORS Ni, J. and Ruben, S. M.  
TITLE Human serine protease and serpin polypeptides  
JOURNAL Patent: US 6566498-A 5 20-MAY-2003;  
FEATURES  
source  
1..840  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 46.4%; Score 506.4; DB 6; Length 840;  
Best Local Similarity 97.9%; Pred. No. 1.1e-104;  
Matches 513; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 568 CACCACGGAACCCATTCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATCGTCTCCCAT 627  
Db 304 CAGCGGAGACCCATTCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATCGTCTCCCAT 363  
QY 628 GCCACCTGCCATGGTGTATCCCGGAGAAATCAGAGCAACATGGTGTGCGAGGCGGC 687  
Db 364 GCCACCTGCCATGGTGTATCCCGGAGAAATCAGAGCAACATGGTGTGCGAGGCGGC 423  
QY 688 GTCCCGGGCAGGATGCTGCCAGGGTGAATCTGGGGGCCCTGTGGGAGAGTGC 747  
Db 424 GTCCCGGGCAGGATGCTGCCAGGGTGAATCTGGGGGCCCTGTGGGAGAGTGC 483  
QY 748 CTTCAAGTCTGGTCTCTGGGGTCTGTGGGGCCCTGTGGCAAGATGGCATCCCTGGA 807  
Db 484 CTTCAAGTCTGGTCTCTGGGGTCTGTGGGGCCCTGTGGCAAGATGGCATCCCTGGA 543  
QY 808 GTCTACACTATATTGCAAGTATGTGACTGGATCCGGATGATCATGAGGAACAATGA 867  
Db 544 GTCTACACTATATTGCAAGTATGTGACTGGATCCGGATGATCATGAGGAACAATGA 603  
QY 868 CTTGTTCTCCACTCCACCCCTTAACCTTGGGTAACTTGGCTCCCTCAGAGC 927  
Db 604 CTTGTTCTCCACTCCACCCCTTAACCTTGGGTAACTTGGCTCCCTCAGAGC 663  
QY 928 ACCAATATCTCTCCATCACTCCCTAGCTCCACTTCTTTGGCTGGGAATCTTTGG 987  
Db 664 ACCAATATCTCTCCATCACTCCCTAGCTCCACTTCTTTGGCTGGGAATCTTTGG 723  
QY 988 AACTTTAACTCTGCGAGCCCTTCTAAGACCCACGAGGGGGTGGAGAAAGTGTCAATA 1047  
Db 724 AACTTTAACTCTGCGAGCCCTTCTAAGACCCACGAGGGGGTGGAGAAAGTGTCAATA 783

QY 1048 GTCTGGAATAAATATAAATGAAGGAGGGGCAAAAAA 1091  
Db 784 GTCTGGAATAAATATAAATGAAGGAGGGGCAAAAAA 827  
RESULT 5  
BD130047  
LOCUS  
DEFINITION Human serine protease and serpin polypeptide.  
ACCESSION BD130047  
VERSION BD130047.1 GI:23224992  
KEYWORDS JP 2002502600-A/3  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 840)  
AUTHORS Ruben, S. M. and Ni, J.  
TITLE Human serine protease and serpin polypeptide  
JOURNAL Patent: JP 2002502600-A 3 29-JAN-2002;  
COMMENT HUMAN GENOME SCIENCES INC  
OS Homo sapiens (human)  
PN JP 2002502600-A/3  
PD 29-JAN-2002  
PF 04-FEB-1999 JP 2000530597  
PR 06-FEB-1998 US 60/073961  
PC STEVEN M RUBEN, JIAN NI  
C12N9/64, A61K38/48, A61P7/02, A61P7/04, A61P7/06, A61P29/00, A61P31/ PC  
00, A61P35/00, A61P37/06, C07K14/81, C07K16/40, C12N1/15, C12N1/19, PC  
C12N1/21  
PC C12N5/10, C12N15/09, C12Q1/37, A61K37/547, C12N5/00, C12N15/00 CC  
Human serine protease and serpin polypeptide  
FH Key Location/Qualifiers  
FT CDS (115)..(603).  
FEATURES  
source  
1..840  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 46.4%; Score 506.4; DB 6; Length 840;  
Best Local Similarity 97.9%; Pred. No. 1.1e-104;  
Matches 513; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 568 CACCACGGAACCCATTCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATCGTCTCCCAT 627  
Db 304 CAGCGGAGACCCATTCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATCGTCTCCCAT 363  
QY 628 GCCACCTGCCATGGTGTATCCCGGAGAAATCAGAGCAACATGGTGTGCGAGGCGGC 687  
Db 364 GCCACCTGCCATGGTGTATCCCGGAGAAATCAGAGCAACATGGTGTGCGAGGCGGC 423  
QY 688 GTCCCGGGCAGGATGCTGCCAGGGTGAATCTGGGGGCCCTGTGGGAGAGTGC 747  
Db 424 GTCCCGGGCAGGATGCTGCCAGGGTGAATCTGGGGGCCCTGTGGGAGAGTGC 483  
QY 748 CTTCAAGTCTGGTCTCTGGGGTCTGTGGGGCCCTGTGGCAAGATGGCATCCCTGGA 807  
Db 484 CTTCAAGTCTGGTCTCTGGGGTCTGTGGGGCCCTGTGGCAAGATGGCATCCCTGGA 543  
QY 808 GTCTACACTATATTGCAAGTATGTGACTGGATCCGGATGATCATGAGGAACAATGA 867  
Db 544 GTCTACACTATATTGCAAGTATGTGACTGGATCCGGATGATCATGAGGAACAATGA 603  
QY 868 CTTGTTCTCCACTCCACCCCTTAACCTTGGGTAACTTGGCTCCCTCAGAGC 927  
Db 604 CTTGTTCTCCACTCCACCCCTTAACCTTGGGTAACTTGGCTCCCTCAGAGC 663  
QY 928 ACCAATATCTCTCCATCACTCCCTAGCTCCACTTCTTTGGCTGGGAATCTTTGG 987  
Db 664 ACCAATATCTCTCCATCACTCCCTAGCTCCACTTCTTTGGCTGGGAATCTTTGG 723  
QY 988 AACTTTAACTCTGCGAGCCCTTCTAAGACCCACGAGGGGGTGGAGAAAGTGTCAATA 1047  
Db 724 AACTTTAACTCTGCGAGCCCTTCTAAGACCCACGAGGGGGTGGAGAAAGTGTCAATA 783

Db 664 ACCAATATCTCTCCATCACTTCCCTAGCTCCACTCTTGTGGCTGGAACTTCTGG 723  
QY 988 AACTTTAACTCTCGCAGCCCTTCTAAGACCCACGACGGGTGAGAGAAGTGTGCAATA 1047  
Db 724 AACTTTAACTCTCGCAGCCCTTCTAAGACCCACGACGGGTGAGAGAAGTGTGCAATA 783  
QY 1048 GTCTGGAATAAATATATAATGAAGGAGGGCCAAAAAAAAAAAAAAA 1091  
Db 784 GTCTGGAATAAATATATAATGAAGGAGGGGAARAAAAAAAAAAAAA 827  
  
RESULT 6  
LOCUS AR263961 619 bp DNA linear PAT 29-JAN-2003  
DEFINITION Sequence 139 from patent US 6331427.  
ACCESSION AR263961  
VERSION AR263961.1 GI:28075965  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 619)  
AUTHORS Robison,K.E.  
TITLE Protease homologs  
JOURNAL Patent: US 6331427-A 139 18-DEC-2001;  
FEATURES Location/Qualifiers  
source 1..619  
/organism="unknown"  
/mol\_type="genomic DNA"  
  
ORIGIN  
Query Match 45.2%; Score 493.2; DB 6; Length 619;  
Best Local Similarity 97.7%; Pred. No. 1.1e-101;  
Matches 543; Conservative 0; Mismatches 8; Indels 5; Gaps 4;  
  
QY 98 CAGACTTTGGAAGTGAACCCACATGGGCTCAGCATCTTTTGTCTCTGTGTTCTTG 157  
Db 67 CAAGACTTTGGAAGTGAACCCACATGGGCTCAGCATCTTTTGTCTCTGTGTTCTTG 126  
QY 158 GACTCAGCCAGCCAGCCACACCGAGATTTCAATGGCACTGAGTGTGGCGTAACCTAC 217  
Db 127 GACTCAGCCAGCCAGCCACACCGAGATTTCAATGGCACTGAGTGTGGCGTAACCTAC 186  
QY 218 AGCCGTGGCAGGTGGGGCTGTTTGAAGGACACGACCTCGCGTGGGGGTGTCCTTATTG 277  
Db 187 AGCCGTGGCAGGTGGGGCTGTTTGAAGGACACGACCTCGCGTGGGGGTGTCCTTATTG 246  
QY 278 ACCAGAGTGGTCTCTCAGCGGCTCCTCAGCGGCACTGACGGGAGCAAGTACTGGGTGGCGCTGG 337  
Db 247 ACCAGAGTGGTCTCTCAGCGGCTCCTCAGCGGCACTGACGGGAGCAAGTACTGGGTGGCGCTGG 306  
QY 338 GGAACACAGCTCAGCAGCTCGACTGACCGAGCAGATCCGACAGCGGCTTCTCTG 397  
Db 307 GGAACACAGCTCAGCAGCTCGACTGACCGAGCAGATCCGACAGCGGCTTCTCTG 366  
QY 398 TGACCCATCCCGCTACCTGGAGCTTCGACGAGCCACGAGCAGCACTCCGCGTCTGC 457  
Db 367 TGACCCATCCCGCTACCTGGAGCTTCGACGAGCCACGAGCAGCACTCCGCGTCTGC 426  
QY 458 GACTGGCTGCGCTGCGGTAAACAGCAGGTTCAACCCCTGCCCTGCCAATGACT 517  
Db 427 GACTGGCTGCGCTGCGGTAAACAGCAGGTTCAACCCCTGCCCTGCCAATGACT 486  
QY 518 GTGCAACCGCTGCGACCGAGTGCAGCTCTCAGGCTGGGGCATCACCAACCCACCGA 577  
Db 487 GTGCAACCGCTGCGACCGAGTGCAGCTCTCAGGCTGGGGCATCACCAACCCACCGA 546  
QY 578 ACCAATCCCGATGCTCCAGTGCTCAACCTCTCCATCTGCTTCCCATGCCACCTGCC 637  
Db 547 ACCAATTC -GGATCTGTCAGTGCCTCAACCTCT -CATGCTCT -CCAATGCCACCTG-C 601  
QY 638 ATGGTGTGTATCCCG 653

Db 602 ATGGTGTGTATCCGG 617  
RESULT 7  
LOCUS AF164623 9120 bp DNA linear PRI 26-JUN-2000  
DEFINITION Homo sapiens trypsin-like serine protease (TLSP) gene, complete cds.  
ACCESSION AF164623  
VERSION AF164623.1 GI:5713130  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 9120)  
AUTHORS Yousef,G.M., Scorilas,A. and Diamandis,E.P.  
TITLE Genomic organization, mapping, tissue expression, and hormonal regulation of trypsin-like serine protease (TLSP PRSS20), a new member of the human kallikrein gene family  
JOURNAL Genomics 63 (1), 88-96 (2000)  
MEDLINE 20130117  
PUBMED 10662548  
REFERENCE 2 (bases 1 to 9120)  
AUTHORS Yousef,G.M., Scorilas,A. and Diamandis,E.P.  
TITLE Direct Submission  
JOURNAL Submitted (01-JUL-1999) Pathology and Laboratory Medicine, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1x5, Canada  
FEATURES Location/Qualifiers  
source 1..9120  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19q13.3-q13.4"  
2313..7622  
/gene="TLSP"  
join(2313..2398,4189..4263,5061..5217,5545..5810,6627..6763,7158..7622)  
/gene="TLSP"  
/product="trypsin-like serine protease"  
join(4224..4263,5061..5217,5545..5810,6627..6763,7158..7310)  
/gene="TLSP"  
/note="kallikrein-like serine protease"  
/codon\_start=1  
/product="trypsin-like serine protease"  
/protein\_id="AAD47815.1"  
/db\_xref="GI:5713131"  
/translation="WRILQLILLALATGLVGGTFLIKGFBECKPHSQWQALFKTRLLCGATLAPRWLLTAHCLPRYIVHLGQHNLRKEGCEQIRRTATESFPHPFNNSL FKNDRNIDMLVKWASPVSTIWAVRPLTSSRCVTAGTSLISGWSGTSFQLRLPHT LFCANITTIHQKCNAYPGNITDTMVCASVQEGGKSCQSDSGGPLVCNLSLQGIIS WQDPCATIRKPGVYTKVKYVDWIQTMMKN"  
  
ORIGIN  
Query Match 34.2%; Score 373.2; DB 9; Length 9120;  
Best Local Similarity 99.2%; Pred. No. 2e-74;  
Matches 375; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 702 TGCCCTCCAGGGTGAATCTCGGGGCCCTCTGGTGTGTGGGGAGTCCCTCAAGGCTGTGT 761  
Db 388 TCCCTCCAGGGTGAATCTCGGGGCCCTCTGGTGTGTGGGGAGTCCCTCAAGGCTGTGT 447  
QY 762 GTCCTGGGGTGTGTGGGGCCCTGTGGCAAGATGGCATCCCTGGAGTCTACACCTATAT 821  
Db 448 GTCCTGGGGTGTGTGGGGCCCTGTGGCAAGATGGCATCCCTGGAGTCTACACCTATAT 507  
QY 822 TTGCAAGTATGTGGACTGGATCCGGATGATCATAGGAACTGACCTGTTTCTCCAC 881  
Db 508 TTGCAAGTATGTGGACTGGATCCGGATGATCATAGGAACTGACCTGTTTCTCCAC 567

```

QY      882  CTCACCCCCACCCCTTAACCTGGGTACCCCTCTGGCCCTCAGACACCAATATCTCTC 941
Db      568  CTCACCCCCACCCCTTAACCTGGGTACCCCTCTGGCCCTCAGACACCAATATCTCTC 627
QY      942  CATCACTTCCCTAGCTCCACTCTTGTGGCCCTGGGAACCTCTTGGAACTTTAACTCTCTG 1001
Db      628  CATCACTTCCCTAGCTCCACTCTTGTGGCCCTGGGAACCTCTTGGAACTTTAACTCTCTG 687
QY      1002 CCAGCCCTTCTAAGACCCACGACGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAAATA 1061
Db      688  CCAGCCCTTCTAAGACCCACGACGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAAATA 747
QY      1062 TAAATGAAGGAGGGGCAA 1079
Db      748  TAAATGAAGGAGGGGCCA 765

RESULT 8
LOCUS   AF135025              11820 bp      DNA      linear      PRI 27-JUN-2000
DEFINITION Homo sapiens kallikrein-like protein 5 gene, alternative splice
ACCESSION AF135025
VERSION   AF135025
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 2909 to 7847)
AUTHORS  Yousef,G.M., Luo,L. and Diamandis,E.P.
TITLE    Identification of novel human kallikrein-like genes on chromosome
          19q13.3-q13.4
JOURNAL  Anticancer Res. 79, 2843-2852 (1999)
REFERENCE 2 (bases 1 to 11820)
AUTHORS  Diamandis,E.P., Yousef,G.M., Luo,L.Y., Magklara,A. and Obiezu,C.V.
TITLE    The new human kallikrein gene family: implications in
          carcinogenesis
JOURNAL  Trends Endocrinol. Metab. 11 (2), 54-60 (2000)
MEDLINE  21121728
PUBMED   10675891
REFERENCE 3 (bases 1 to 11820)
AUTHORS  Yousef,G.M., Magklara,A. and Diamandis,E.P.
TITLE    KLK-L5 is a novel serine protease and a new member of the human
          kallikrein gene family-differential expression in breast cancer
          unpublished
JOURNAL  4 (bases 1 to 11820)
REFERENCE 4 (bases 1 to 11820)
AUTHORS  Yousef,G.M., Magklara,A., Scorilas,A. and Diamandis,E.P.
TITLE    Cloning of new alternatively spliced forms of the kallikrein-like
          gene 5 (KLK-L5)
JOURNAL  Unpublished
REFERENCE 5 (bases 2909 to 7847)
AUTHORS  Yousef,G.M., Magklara,A. and Diamandis,E.P.
TITLE    Direct Submission
JOURNAL  Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount
          Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
          Canada
REFERENCE 6 (bases 1 to 11820)
AUTHORS  Yousef,G.M., Magklara,A. and Diamandis,E.P.
TITLE    Direct Submission
JOURNAL  Submitted (01-NOV-1999) Pathology and Laboratory Medicine, Mount
          Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
          Canada
REMARK   Sequence update by submitter
COMMENT  On Nov 1, 1999 this sequence version replaced gi:4589276.
FEATURES             location/Qualifiers
                     1..11820
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /map="19q13.3-q13.4"

```

```

mRNA
join(3743..3840,3995..4050,4496..4655,6500..6759,
7714..7847,9178..9543)
/product="kallikrein-like protein 5"
join(3743..3840,3995..4050,4496..4655,6500..6759,
7714..7847,9178..9293,9423..9543)
/product="kallikrein-like protein 5-related protein 1"
join(3743..3840,3995..4050,4496..4655,7714..7847,
9178..9543)
/product="kallikrein-like protein 5-related protein 2"
join(4014..4050,4496..4655,6500..6759,7714..7847,
9178..9293,9423..9480)
/notes="KLK-L5-related protein 1; kallikrein-like serine
protease; alternatively spliced"
/codon_start=1
/product="kallikrein-like protein 5-related protein 1"
/protein_id="AAF06065.1"
/db_xref="GI:6249632"
/translation="MGLSIFLLLCVLGLSOAATPKIENGTECGRNSOPWQVGLPEGNS
LRGGVLIDHRWVLTAAHCSGRYVWRLGHSLSQLDWTQIRHSGFSVTHPGVLGAS
TSHEHDLRLRLPLPVRTSVQPLPNDCAATAGTCHVSGWGTINHPNPFDDLQ
CLNLSVSHATCHGVYPRITSNMVCAGVPGQDACQDSGGPLVCGVLQGLVSWGS
VPGCGQIGPVYTYICNLTGLGTSWNFNSQPF"
join(4014..4050,4496..4655,6500..6759,7714..7847,
9178..9333)
/notes="KLK-L5"
/codon_start=1
/product="kallikrein-like protein 5"
/protein_id="AAD26426.2"
/db_xref="GI:6166249"
/translation="MGLSIFLLLCVLGLSOAATPKIENGTECGRNSOPWQVGLPEGNS
LRGGVLIDHRWVLTAAHCSGRYVWRLGHSLSQLDWTQIRHSGFSVTHPGVLGAS
TSHEHDLRLRLPLPVRTSVQPLPNDCAATAGTCHVSGWGTINHPNPFDDLQ
CLNLSVSHATCHGVYPRITSNMVCAGVPGQDACQDSGGPLVCGVLQGLVSWGS
VPGCGQIGPVYTYICNLTGLGTSWNFNSQPF"
join(4014..4050,4496..4655,7714..7847,9178..9182)
/notes="KLK-L5-related protein 2; kallikrein-like serine
protease; alternatively spliced"
/codon_start=1
/product="kallikrein-like protein 5-related protein 2"
/protein_id="AAF06066.1"
/db_xref="GI:6249633"
/translation="MGLSIFLLLCVLGLSOAATPKIENGTECGRNSOPWQVGLPEGNS
LRGGVLIDHRWVLTAAHCSGRIPGPSAPVPLRLPCHLPWCVSRENHEQHVCR
RPGAGCLPG"
polyA_signal 9520..9525
ORIGIN
Query Match      34.2%; Score 373.2; DB 9; Length 11820;
Best Local Similarity 99.2%; Pred. No. 2e-74;
Matches 375; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY      702  TGCCTGCCAGGGTGATTCCTGGGGCCCCCTGGTGTGTGGGGGAGTCCTTCAAGTCTGGT 761
Db      9168  TCCCTTCAGGGTGATTCCTGGGGCCCCCTGGTGTGTGGGGGAGTCCTTCAAGTCTGGT 9227
QY      762  GTCCTGGGGGCTCTGGGGCCCCCTGGCACAAGATGGCATCCCTGGAGTCTACACCTATAT 821
Db      9228  GTCTCTGGGGGCTCTGGGGCCCCCTGTGGACAGATGGCATCCCTGGAGTCTACACCTATAT 9287
QY      822  TTGCAAGTATGTGAGTCGGATCCGGATGATCATGAGGAACAACAGCTGTCTTCTCCAC 881
Db      9288  TTGCAAGTATGTGAGTCGGATCCGGATGATCATGAGGAACAACAGCTGTCTTCTCCAC 9347
QY      882  CTCCACCCCCACCCCTTAACCTGGGTACCCCTCTGGGCCCTCAGAGCAACCAATATCTCTC 941
Db      9348  CTCCACCCCCACCCCTTAACCTGGGTACCCCTCTGGGCCCTCAGAGCAACCAATATCTCTC 9407
QY      942  CATCACTTCCCTAGCTCCACTCTTGTGGCTGGGAACCTCTTGGAACTTTAACTCTCTG 1001
Db      9408  CATCACTTCCCTAGCTCCACTCTTGTGGCTGGGAACCTCTTGGAACTTTAACTCTCTG 9467
QY      1002 CCAGCCCTTCTAAGACCCACGACGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAAATA 1061

```





	Query Match	34.2%;	Score 373.2;	DB 9;	Length 132323;
	Best Local Similarity	99.2%;	Pred. No. 1.8e-74;		
	Matches 375;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	702	TGCTCTCCAGGGTGATCTCGGGGGCCCCCTGTGTGTGTGGGGAGTCCTTCAAGGTCCTGGT	761		
Db	29548	TCCCTTCCAGGGTGATCTCGGGGGCCCCCTGTGTGTGTGGGGAGTCCTTCAAGGTCCTGGT	29489		
QY	762	GTCTCTGGGGGTCTGTGGGGCCCTGTGGACAAGATGCATCCCTGGAGTGCTACACTATAT	821		



```

/protein_id="AAG33362.1"
/db_xref="GI:11244767"
/translation="MKLGLLCAALLSLAGHWADTRAIGAEBCRPNQSPQWQAGLFLHIT
RLFCGATLIISDRMLLTAHCRKPYLWRLGHEHLHWEGPEQLFRVTDFFHPGFENKD
LSANDHNDIMLRKLPQARLSPAVQPLNLSOTCVSPMOCLISGWAQVSPKALFV
TLOCANTISILENKLCHWAYPCHISDSMLCAGLWEGGRGSCQDGGFLVCNGTLAGVV
SGNAPCSRPRPRPVYTSVCHYLDWIOEIMEN"

Query Match      34.2%; Score 373.2; DB 9; Length 230000;
Best Local Similarity 99.2%; Pred. No. 1.7e-74;
Matches 375; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 702 TGCTGTCGACGGTGTATTCTGGGGGCCCTCTGCTGTGTGGGGAGTCTCTCAAGGCTCTGGT 761
Db 217331 TCCCTTCAGAGGTGATCTCTGGGGGCCCTCTGCTGTGTGGGGAGTCTCTCAAGGCTCTGGT 217272
QY 762 GTCTCGGGGCTCTGTGGGGCCCTCTGGACAAGATGGATCCCTCGAGTCTACACCTATAT 821
Db 217271 GTCTCGGGGCTCTGTGGGGCCCTCTGGACAAGATGGATCCCTCGAGTCTACACCTATAT 217212
QY 822 TTGCAAGTATGTGACTGGATCCGATGATCATGAGNAACAAGTCTTTCTCTCCAC 881
Db 217211 TTGCAAGTATGTGACTGGATCCGATGATCATGAGNAACAAGTCTTTCTCTCCAC 217152
QY 882 CTCCACCCCAACCCCTTTAACTTGGGTACCCCTCTGGCCCTCAGAGCAACAATATCTCTC 941
Db 217151 CTCCACCCCAACCCCTTTAACTTGGGTACCCCTCTGGCCCTCAGAGCAACAATATCTCTC 217092
QY 942 CATCACTTCCCTAGTCCCACTCTTGTGGCTCGGAACTCTTTGGAACCTTAACTCTG 1001
Db 217091 CATCACTTCCCTAGTCCCACTCTTGTGGCTCGGAACTCTTTGGAACCTTAACTCTG 217032
QY 1002 CCAGCCCTTCAAGACCCACAGCGGGGTGAGAGAGTGTCAATAGTCTGGAATAATA 1061
Db 217031 CCAGCCCTTCAAGACCCACAGCGGGGTGAGAGAGTGTCAATAGTCTGGAATAATA 216972
QY 1062 TAAATGAAGGAGGGGCA 1079
Db 216971 TAAATGAAGGAGGGGCA 216954

```

```

RESULT 11
AC140096/c
LOCUS      AC140096      174724 bp      DNA      linear      HTG 14-JUN-2003
DEFINITION Pan troglodytes clone RP43-63K10, WORKING DRAFT SEQUENCE, 10
ordered pieces.
ACCESSION  AC140096
VERSION    AC140096.4  GI:31745211
KEYWORDS   HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 174724)
Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G., Brinkley,C., Brooks,S., Cariaga,K.,
Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
Hurle,B., Idol,J.R., Karlins,E., Kwong,P., Lalic,P., Lee-Il-in,S.-Q.,
Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
Maskeri,B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E.,
Prasad,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
Sison,C., Stantropop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 174724)
Green,E.D.
Direct Submission
TITLE      Direct Submission
JOURNAL    Submitted (21-FEB-2003) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 174724)
Green,E.D.

```

# TITLE JOURNAL COMMENT

Direct Submission  
Submitted (14-JUN-2003) NIH Intramural Sequencing Center, 8717  
Government Circle, Gaithersburg, MD 20877, USA  
On Jun 14, 2003 this sequence version replaced gi:29725783.  
----- Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: http://www.nisc.nih.gov  
Contact: nisc.zoonhgri.nih.gov  
----- Project Information  
Center project name: dxz  
Center clone name: 063K10

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 172841 bases at least Q40  
Consensus quality: 173490 bases at least Q30  
Consensus quality: 173737 bases at least Q20  
Insert size: 150000; agarose-fp  
Quality coverage: 15.91x in Q20 bases; agarose-fp  
Quality coverage: 13.73x in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and the accession number will be preserved.

1 47270: contig of 47270 bp in length  
\* 47271 47370: gap of unknown length  
\* 47371 58115: contig of 10745 bp in length  
\* 58116 58215: gap of unknown length  
\* 58216 61050: contig of 2835 bp in length  
\* 61051 61150: gap of unknown length  
\* 61151 65346: contig of 4196 bp in length  
\* 65347 71536: gap of unknown length  
\* 71537 71636: contig of 6090 bp in length  
\* 71637 83139: gap of unknown length  
\* 83140 83239: gap of unknown length  
\* 83240 85133: contig of 1894 bp in length  
\* 85134 85233: gap of unknown length  
\* 85234 88937: contig of 3604 bp in length  
\* 88938 88937: gap of unknown length  
\* 88938 129108: contig of 40171 bp in length  
\* 129109 129209: gap of unknown length  
\* 129209 174724: contig of 45516 bp in length.

FEATURES  
Location/Qualifiers  
1. .174724  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone\_lib="RP43-63K10"  
/clone\_lib="RP43"  
misc\_feature 1. .79098

```
/notes="clone overlaps with GenBank Accession Number
AC130782 clone CH251-355A20 (center project name dhz)"
misc_feature 1..47270
/notes="assembly_fragment
clone_end:T7
vector_side:left"
misc_feature 47371..58115
/notes="assembly_fragment"
misc_feature 58216..61050
/notes="assembly_fragment"
misc_feature 61151..65346
/notes="assembly_fragment"
misc_feature 65447..71536
/notes="assembly_fragment"
misc_feature 71637..83139
/notes="assembly_fragment"
misc_feature 83240..85133
/notes="assembly_fragment"
misc_feature 85234..88837
/notes="assembly_fragment"
misc_feature 88938..129108
/notes="assembly_fragment"
misc_feature 113257..174724
/notes="clone overlaps with GenBank Accession Number
AC133392 clone CH251-126O24 (center project name dxa)"
129209..174724
/notes="assembly_fragment
clone_end:SP6
vector_side:right"

ORIGIN
Query Match 33.8%; Score 368.4; DB 2; Length 174724;
Best Local Similarity 98.4%; Pred. No. 2.2e-73;
Matches 372; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 702 TGCTCCGAGGGTGATTCTGGGGCCCCCTGGTGTGTGGGGAGTCTTCAAGTCTGGT 761
Db 38231 TCCCTTCCAGGGTGATTCTGGGGCCCCCTGGTGTGTGGGGAGTCTTCAAGTCTGGT 38172
QY 762 GTCCTGGGGCTCTGGGGCCCTGTGGACAAGATGTCATCCCTGGAGTCTACACCTATAT 821
Db 38171 GTCCTGGGGCTCTGGGGCCCTGTGGACAAGATGTCATCCCTGGAGTCTACACCTATAT 38112
QY 822 TTGCAAGTATGTGAGTGGATCCGATGATCATGAGGAACTGCTTTTCCCTCCAC 881
Db 38111 TTGCAAGTATGTGAGTGGATCCGATGATCATGAGGAACTGCTTTTCCCTCCAC 38052
QY 882 CTCACCCCCACCCCTTAAGTGGGTACCCCTCTGGCCCTCAGACCAATATCTCCTC 941
Db 38051 CTCACCCCCACCCCTTAAGTGGGTACCCCTCTGGCCCTCAGACCAATATCTCCTC 37992
QY 942 CATCACTTCCCTTAGCTCCACTCTGTGGCCCTGGGAATCTTTGGAATTTAACTCCTG 1001
Db 37991 CATCACTTCCCTTAGCTCCACTCTGTGGCCCTGGGAATCTTTGGAATTTAACTCCTG 37932
QY 1002 CCAGCCCTTCTAAGACCCAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAAATAATA 1061
Db 37931 CCAGCCCTTCTAAGACCCAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAAATAATA 37872
QY 1062 TAAATGAGGGGGGCAA 1079
Db 37871 TAAATGAGGGGGGCAA 37854

RESULT 12
AC130782/c
LOCUS AC130782 200792 bp DNA linear HTG 22-NOV-2002
DEFINITION Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12
ordered pieces.
ACCESSION AC130782
VERSION AC130782.2 GI:25167101
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Pan troglodytes (Chimpanzee)
```

## ORGANISM

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE  
AUTHORS

1. (bases 1 to 200792)

Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,  
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,  
Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,  
Haghghi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,  
Laric,P., Lee-Lin,S.-O., Legaspi,R., Maskeri,B., McDowell,J.,  
Margulies,E.H., Masiello,C., Pearson,R., Portnoy,M.B., Prasad,A.,  
Paguirigan,C., Pearson,R., Portnoy,M.B., Prasad,A.,  
Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,  
Stantipoy,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,  
Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.

NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 200792)

Green,E.D.

Direct Submission

Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717  
Grovemont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 200792)

Green,E.D.

Direct Submission

Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717  
Grovemont Circle, Gaithersburg, MD 20877, USA

On Nov 22, 2002 this sequence version replaced gi:22218452.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>Contact: [nisc.zoo@nhgri.nih.gov](mailto:nisc.zoo@nhgri.nih.gov)

----- Project Information

Center project name: dhz

Center clone name: 355A20

The sequence data in this record represents an 'enhanced'  
version of a Phase 2 submission. Specifically, the indicated  
order and orientation of each sequence contig has been  
established using one or more of the following: read-pair  
data from individual subclones, overlaps with neighboring  
clones, alignment with available reference sequence (e.g.,  
human), and/or confirmation by PCR testing. In addition,  
the sequence assembly is based on at least 8X average  
coverage in Q20 bases and has been reviewed to rule out  
gross misassemblies, the low-quality ends of sequence  
contigs have been trimmed away, and each base is associated  
with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 196556 bases at least Q40

Consensus quality: 197883 bases at least Q30

Consensus quality: 198879 bases at least Q20

Insert size: 215000; agarose-fp

Quality coverage: 9.03x in Q20 bases; agarose-fp

Quality coverage: 9.73x in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 28306: contig of 28306 bp in length  
\* 28307 28406: gap of unknown length  
\* 28407 37856: contig of 9450 bp in length  
\* 37857 37956: gap of unknown length

```

* 37957 73522: contig of 35566 bp in length
* 73523 73622: gap of unknown length
* 73623 83567: contig of 9945 bp in length
* 83568 83667: gap of unknown length
* 83668 8817: contig of 5150 bp in length
* 8818 88917: gap of unknown length
* 88918 125611: contig of 36694 bp in length
* 125612 125711: gap of unknown length
* 125712 159879: contig of 34168 bp in length
* 159880 159979: gap of unknown length
* 159980 174698: contig of 14719 bp in length
* 174699 174798: gap of unknown length
* 174799 186382: contig of 11584 bp in length
* 186383 186482: gap of unknown length
* 186483 193444: contig of 6862 bp in length
* 193445 193444: gap of unknown length
* 193445 199363: contig of 5919 bp in length
* 199364 199463: gap of unknown length
* 199464 200792: contig of 1329 bp in length.

```

## FEATURES

```

source
1..200792
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="CH251-355A20"
/clone_lib="CH251"

```

## misc\_feature

```

1..28306
/note="assembly_fragment"
clone_end:SP6
vector_side:left

```

## misc\_feature

```

28407..37856
/note="assembly_fragment"

```

## misc\_feature

```

37957..73522
/note="assembly_fragment"

```

## misc\_feature

```

73623..83567
/note="assembly_fragment"

```

## misc\_feature

```

83668..88817
/note="assembly_fragment"

```

## misc\_feature

```

88918..125611
/note="assembly_fragment"

```

## misc\_feature

```

125712..159879
/note="assembly_fragment"

```

## misc\_feature

```

159980..174698
/note="assembly_fragment"

```

## misc\_feature

```

174799..186382
/note="assembly_fragment"

```

## misc\_feature

```

186483..193344
/note="assembly_fragment"

```

## misc\_feature

```

193445..199363
/note="assembly_fragment"

```

## misc\_feature

```

199464..200792
/note="assembly_fragment"
clone_end:17
vector_side:right

```

## ORIGIN

```

Query Match      33.8%; Score 368.4; DB 2; Length 200792;
Best Local Similarity 98.4%; Pred. No. 2.e-73;
Matches 372; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 702 TGCCTGCCAGGTGATTCTGGGGGCCCCCTGGTGTGTGGGGAGTCTTCAAGGTCCTGT 761
DB 148663 TCCCTTCCAGGGTGATTCTGGGGGCCCCCTGGTGTGTGGGGAGTCCITCAAGGTCCTGT 148604

QY 762 GTCCTGGGGGCTCTGGGGGCCCCCTGGACAGATGGCATCCCTGGAGTCTACACCTATAT 821
DB 148603 GTCCTGGGGGCTCTGGGGGCCCCCTGGACAGATGGCATCCCTGGAGTCTACACCTATAT 148544

QY 822 TTGCAAGTATGTGGACTGGATCCGGATGATCATGAGGAACAACACTGACCTGTTCTCTCCAC 881
DB 148543 TTGCAAGTATGTGGACTGGATCCGGATGATCATGAGGAACAACACTGACCTGTTCTCTCCAC 148484

QY 882 CTCACACCCCAACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGACACAATATCTCTCTC 941

```

```

Db 148483 CTCACACCCCAACCCCTTAGCTTGGGTACCACTCTGGCCCCCAGAGACCAATATCTCTC 148424
QY 942 CATCACTTCCCTTAGCTCCCACTCTTGTGGCTTGGGAACCTTCTTGGAACTTTAACTCTG 1001
Db 148423 CATCACTTCCCTTAGCTCCCACTCTTGTGGCTTGGGAACCTTCTTGGAACTTTAACTCTG 148364
QY 1002 CGAGCCCTTAAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAAATAAATA 1061
Db 148363 CGAGCCCTTAAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAAATAAATA 148304
QY 1062 TAAATGAAGGAGGGCAA 1079
Db 148303 TAAATGAAGGAGGGCAA 148286

RESULT 13
AC073185/c
LOCUS AC073185 142334 bp DNA linear HTG 17-AUG-2000
DEFINITION Homo sapiens chromosome 19 clone RP11-10111, WORKING DRAFT
SEQUENCE, 32 unordered pieces.
ACCESSION AC073185
VERSION AC073185.4 GI:9838034
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 142334)
AUTHORS Waterston,R.H.
JOURNAL Unpublished
TITLE The sequence of Homo sapiens clone
REFERENCE 2 (bases 1 to 142334)
AUTHORS Waterston,R.H.
JOURNAL Direct Submission
TITLE Submitted (10-JUN-2000) Genome Sequencing Center, Washington
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Aug 17, 2000 this sequence version replaced gi:9653152.

```

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH001011
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: bases at least Q40
Consensus quality: bases at least Q30
Consensus quality: bases at least Q20
Insert size: 135000; agarose-fp
Insert size: 139234; sum-of-contigs
Quality coverage: 3.86 in Q20 bases; agarose-fp
Quality coverage: 3.87 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1452: contig of 1452 bp in length
* 1453 1552: gap of unknown length
* 1553 2737: contig of 1185 bp in length
* 2738 2837: gap of unknown length
* 2838 4665: contig of 1828 bp in length

```

```
* 4666 4665: gap of unknown length
* 4766 6054: contig of 1289 bp in length
* 6055 6154: gap of unknown length
* 6155 7791: contig of 1637 bp in length
* 7792 7891: gap of unknown length
* 7892 10026: contig of 2135 bp in length
* 10027 10126: gap of unknown length
* 10127 11895: contig of 1769 bp in length
* 11896 11995: gap of unknown length
* 11996 15070: contig of 3075 bp in length
* 15071 15170: gap of unknown length
* 15171 18476: contig of 3306 bp in length
* 18477 18576: gap of unknown length
* 18577 21537: contig of 2961 bp in length
* 21538 21637: gap of unknown length
* 21638 24625: contig of 2988 bp in length
* 24626 24726: gap of unknown length
* 24727 27583: contig of 2858 bp in length
* 27584 27683: gap of unknown length
* 27684 30892: contig of 3209 bp in length
* 30893 30992: gap of unknown length
* 30993 35210: contig of 4217 bp in length
* 35211 35309: gap of unknown length
* 35310 39261: contig of 3952 bp in length
* 39262 39361: gap of unknown length
* 39362 42356: contig of 2995 bp in length
* 42357 42456: gap of unknown length
* 42457 46421: contig of 3965 bp in length
* 46422 46521: gap of unknown length
* 46522 51157: contig of 4636 bp in length
* 51158 51257: gap of unknown length
* 51258 57416: contig of 6159 bp in length
* 57417 57516: gap of unknown length
* 57517 61710: contig of 4194 bp in length
* 61711 61810: gap of unknown length
* 61811 66249: contig of 4439 bp in length
* 66250 66349: gap of unknown length
* 66350 71980: contig of 5631 bp in length
* 71981 72080: gap of unknown length
* 72081 77454: contig of 5374 bp in length
* 77455 77554: gap of unknown length
* 77555 84318: contig of 6764 bp in length
* 84319 84418: gap of unknown length
* 84419 88649: contig of 4231 bp in length
* 88650 88749: gap of unknown length
* 88750 94419: contig of 5670 bp in length
* 94420 94519: gap of unknown length
* 94520 100165: contig of 5846 bp in length
* 100166 100265: gap of unknown length
* 100266 107334: contig of 7069 bp in length
* 107335 107434: gap of unknown length
* 107435 115093: contig of 7659 bp in length
* 115094 115193: gap of unknown length
* 115194 122023: contig of 6830 bp in length
* 122024 122123: gap of unknown length
* 122124 132649: contig of 10526 bp in length
* 132650 132749: gap of unknown length
* 132750 142334: contig of 9585 bp in length.
```

## FEATURES

source

```
1..142334
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="RP11-10111"
1..1452
/note="assembly_name:Contig26"
1553..2737
/note="assembly_name:Contig27"
2838..4665
/note="assembly_name:Contig28"
4766..6054
/note="assembly_name:Contig29"
```

```
misc_feature 6155..7791
/note="assembly_name:Contig30"
misc_feature 7892..10026
/note="assembly_name:Contig31"
misc_feature 10127..11895
/note="assembly_name:Contig33"
misc_feature 11996..15070
/note="assembly_name:Contig34"
misc_feature 15171..18476
/note="assembly_name:Contig35"
misc_feature 18577..21537
/note="assembly_name:Contig36"
misc_feature 21638..24625
/note="assembly_name:Contig37"
misc_feature 24726..27583
/note="assembly_name:Contig38"
misc_feature 27684..30892
/note="assembly_name:Contig39"
misc_feature 30993..35209
/note="assembly_name:Contig40"
misc_feature 35310..39261
/note="assembly_name:Contig41"
misc_feature 39362..42356
/note="assembly_name:Contig42"
misc_feature 42457..46421
/note="assembly_name:Contig43"
misc_feature 46522..51157
/note="assembly_name:Contig44"
misc_feature 51258..57416
/note="assembly_name:Contig45"
misc_feature 57517..61710
/note="assembly_name:Contig46"
misc_feature 61811..66249
/note="assembly_name:Contig47"
misc_feature 66350..71980
/note="assembly_name:Contig48"
misc_feature 72081..77454
/note="assembly_name:Contig49"
misc_feature 77555..84318
/note="assembly_name:Contig50"
misc_feature 84419..88649
/note="assembly_name:Contig51"
misc_feature 88750..94419
/note="assembly_name:Contig52"
misc_feature 94520..100165
/note="assembly_name:Contig53"
misc_feature 100266..107334
/note="assembly_name:Contig54"
misc_feature 107435..115093
/note="assembly_name:Contig55"
misc_feature 115194..122023
/note="assembly_name:Contig56"
misc_feature 122124..132649
/note="assembly_name:Contig57"
clone_end:17
vector_side:right"
132750..142334
/note="assembly_name:Contig58"
```

## ORIGIN

```
Query Match 33.1%; Score 361.2; DB 2; Length 142334;
Best Local Similarity 98.9%; Pred. No. 9.6e-72;
Matches 374; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 702 TGCTGCGAGGTGATTCTGGGGGCCCCCTGGTGTGTGGGGAGTCCCTCAAGTCTGGT 761
Db 10990 TCCCTTCAGGGTATTCTGGGGGCCCCCTGGTGTGTGGGGAGTCCCTCAAGTCTGGT 10931

Qy 762 GTCTGCGGGGTCTGTGGGGGCCCTGTGGACAAGATGGGATCCCTGAGTCTACACCTATAT 821
Db 10930 GTCTGCGGGGTCTGTGGGGGCCCTGTGGACAAGATGGGATCCCTGAGTCTACACCTATAT 10871

Qy 822 TTGCAAGTATGTGAGCTGGATCCGGATCATATGAGGAAACAATGACCTGTTTCCTCCAC 881
```

Mon Jun 21 09:00:25 2004

Db	10870	TTGCAAGTATGTGACATCGGATCATATGATGAGGAACAACTGACCTGTTTCTCCAC	10811
Qy	882	CTCCACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGAGCACCATAATCTCTC	941
Db	10810	CTCCACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGAGCACCATAATCTCTC	10752
Qy	942	CATCACTTCCCTAGCTCCACTTGTGTGGCTGGGAATCTTGGAACTTTAACTCTG	1001
Db	10751	CATCACTTCCCTAGCTCCACTTGTGTGGCTGGGAATCTTGGAACTTTAACTCTG	10692
Qy	1002	CCAGCCCTTAAAGCCACGAGCGGGTGAGAGAGTGCAGTAAGTCTGCAATAAATA	1061
Db	10691	CCAGCCCTTAAAGCCACGAGCGGGTGAGAGAGTGCAGTAAGTCTGCAATAAATA	10632
Qy	1062	TAAATGAAGAGGGGCA 1079	
Db	10631	TAAATGAAGAGGGGCA 10614	
RESULT 14			
AX070938			
LOCUS	405 bp	DNA	linear
DEFINITION	Sequence 1410 from Patent WO0102568.		PAT 25-JAN-2001
ACCESSION	AX070938		
VERSION	AX070938.1	GI:12581211	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
1	Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J., Kassam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D., Lamson, G., Drmanac, R., Crkenjakov, R., Drmanac, S., Dickson, M., Labat, I., Leshkowitz, D., Kita, D., Garcia, V. and Strache-Crain, B.		
TITLE	Human genes and gene expression products		
JOURNAL	Patent: WO 0102568-A 1410 11-JAN-2001;		
FEATURES	CHIRON CORPORATION (US); HYSEQ, INC. (US)		
source	Location/Qualifiers		
1..405	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	27.2%; Score 297; DB 6; Length 405;		
Best Local Similarity	99.7%; Pred. No. 5.2e-57;		
Matches	308; Conservative 0; Mismatches 0; Indels 1; Gaps 1;		
Qy	10	CATCCCTTGGTACCTCAAGAGAGCAGAGAGGCGAGAGTGGGGGCACAGGAA	69
Db	9	CATCCCTTGGTACCTCAAGAGAGCAGAGAGGCGAGAGTGGGGGCACAGGAA	68
Qy	70	GGGTGACCTCTGAGATCCCTTTCCTCCAGAGCTTTGGAAGTGACCCACATGGGCT	128
Db	69	GGGTGACCTCTGAGATCCCTTTCCTCCAGAGCTTTGGAAGTGACCCACATGGGCT	128
Qy	129	CAGCATCTTTTGTCTGTGTCTTCTTGGGCTGAGCCAGGAGCCACCGAGATTTT	188
Db	129	CAGCATCTTTTGTCTGTGTCTTCTTGGGCTGAGCCAGGAGCCACCGAGATTTT	188
Qy	189	CAATGGCACTGAGTGGCGGTAATCTCAGCCGCGGAGGCTGTTTGGGGCAC	248
Db	189	CAATGGCACTGAGTGGCGGTAATCTCAGCCGCGGAGGCTGTTTGGGGCAC	248
Qy	249	CAGCCTGGGCTCGGGGGTGTCTTATTGACACAGAGTGGGTCTCAGAGGGCTCACTG	308
Db	249	CAGCCTGGGCTCGGGGGTGTCTTATTGACACAGAGTGGGTCTCAGAGGGCTCACTG	308
Qy	309	CAGCGGCG 317	
Db	309	CAGCGGCG 317	

RESULT 15  
AC130188/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AC130188 176647 bp DNA linear HTG 14-NOV-2002  
Papio anubis clone RP41-421P3, WORKING DRAFT SEQUENCE, 12 ordered pieces.  
AC130188  
AC130188.2 GI:24960890  
HTG; HTGS\_PHASE2; HTGS\_DRAFT.  
Papio anubis (olive baboon)  
Papio anubis  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Papio.  
1 (bases 1 to 176647)  
Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laric, P., Lee-Lin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masiello, C., Maskeri, B., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C., Stantripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 176647)  
Green, E.D.  
Direct Submission  
Submitted (08-AUG-2002) NIH Intramural Sequencing Center, 8717  
Grovemont Circle, Gaithersburg, MD 20877, USA  
3 (bases 1 to 176647)  
Green, E.D.  
Direct Submission  
Submitted (14-NOV-2002) NIH Intramural Sequencing Center, 8717  
Grovemont Circle, Gaithersburg, MD 20877, USA  
On Nov 14, 2002 this sequence version replaced gi:22138439.  
----- Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: <http://www.nisc.nih.gov>  
Contact: [nisc.zoo@nhgri.nih.gov](mailto:nisc.zoo@nhgri.nih.gov)  
----- Project Information  
Center project name: dsr  
Center clone name: 421P03

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.  
----- Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 171127 bases at least Q40  
Consensus quality: 173424 bases at least Q30  
Consensus quality: 174803 bases at least Q20  
Insert size: 188000; agarose-fp  
Insert size: 175547; sum-of-contigs  
Quality coverage: 8.76x in Q20 bases; agarose-fp  
Quality coverage: 9.39x in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently



\* consists of 12 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.

\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 6526: contig of 6526 bp in length  
\* 6527 6626: gap of unknown length  
\* 18812: contig of 12186 bp in length  
\* 18912: gap of unknown length  
\* 61780: contig of 42868 bp in length  
\* 61880: gap of unknown length  
\* 61881 97068: contig of 35188 bp in length  
\* 97069 97168: gap of unknown length  
\* 125206: contig of 28038 bp in length  
\* 125306: gap of unknown length  
\* 125207 127021: contig of 1715 bp in length  
\* 127022 127121: gap of unknown length  
\* 129387: contig of 2266 bp in length  
\* 129487: gap of unknown length  
\* 132302: contig of 2815 bp in length  
\* 132402: gap of unknown length  
\* 135735: contig of 3333 bp in length  
\* 135835: gap of unknown length  
\* 143558: contig of 7723 bp in length  
\* 143559 143658: gap of unknown length  
\* 170182: contig of 26524 bp in length  
\* 170183 170282: gap of unknown length  
\* 170283 176647: contig of 6365 bp in length.

FEATURES

source

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

ORIGIN

Query Match 26.9%; Score 294; DB 2; Length 176647;  
Best Local Similarity 91.3%; Pred. No. 1.9e-56;

	Matches	347;	Conservative	0;	Mismatches	25;	Indels	8;	Gaps	3;
QY	702	TGCTGTCACGGGTGATTCTCGGGGCCCCCTGGTGTGTGGGGGAGTCCCTTCAAGGTCTGGT	761							
Db	72787	TCCTTTCCAGGGTGATTCTCGGGGCCCCCTGGTGTGTGGGGGAGTCCCTTCAAGGTCTGGT	72728							
QY	762	GTCTGGGGGTCTGTGGGGGCCCTGTGGACAAGATGGCATCCCTGGAGTCTACACCTATAT	821							
Db	72727	GTCTGGGGGTCTGTGGGGGCCCTGTGGACAAGATGGCATCCCTGGAGTCTACACCTATGT	72668							
QY	822	TTGCAAGTATGTGGACTGGATCCGGATGATCATGAGGAACAACCTGCTTTCTCCAC	881							
Db	72667	TTGCAAGTATGTGGACTGGATCCGGATGATCATGAGGAACAACCTGCTTTCTCCAC	72608							
QY	882	CTCCACCCCCACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGAGCACCAATATCTCTCTC	941							
Db	72607	CTCCATCC-----CTTAGCTTGGGTACCCCTCTGGCCCTCAGAGCACCAATATCTCTCTC	72554							
QY	942	CATCACTTCCCTAGCTCCACTCTTGTGGGCTGGAACTTTCTTGGAAC-TTTAACTCCT	1000							
Db	72553	CATCGCTTTTCTGGCTCCATTCTTGTGGGCTGGAACTTTTGGAACTTTTAACTCCT	72494							
QY	1001	GCAGCCCTTCTAAGACCCACGACG-GGGGTGAGAGAGTGTGCAATAGTCTGGAATAAA	1059							
Db	72493	GCAGCCCTTCTAAGACCCACGACGAGCGAGGACGAGAGAAATGTGCAATAGTCTGGAATAAA	72434							
QY	1060	TATTAATGAAGAGGGGCAA	1079							
Db	72433	TACAAATGAAGAGGGGCAA	72414							

Search completed: June 20, 2004, 03:46:14  
Job time : 4628 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2004, 23:53:03 ; Search time 508 Seconds  
(without alignments)  
9123.591 Million cell updates/sec

Title: US-10-015-385A-193

Perfect score: 1091

Sequence: 1 caagcaggtcatcccttgg.....aggggcaaaaaaaaaa 1091

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 29Jan04:\*

- 1: geneseqn1980a:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002s:\*
- 7: geneseqn2003as:\*
- 8: geneseqn2003bs:\*
- 9: geneseqn2003cs:\*
- 10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1091	100.0	1091	3	Aaa77671 Human PRO
2	1091	100.0	1091	3	Aac58114 Human PRO
3	1091	100.0	1091	3	Aaa37075 Human PRO
4	1091	100.0	1091	3	Aaf54341 DNA encod
5	1091	100.0	1091	8	AcD68378 Novel hum
6	1091	100.0	1091	8	AcD68378 Human cdn
7	1091	100.0	1091	8	AcD68024 Novel hum
8	1091	100.0	1091	9	AdC18062 Human PRO
9	1091	100.0	1091	9	AdD70708 Human CDN
10	1091	100.0	1091	9	AdD39785 Human CDN
11	1091	100.0	1091	9	AdD70231 Human CDN
12	1091	100.0	1091	9	AdD38352 Human CDN
13	1091	100.0	1091	9	AdD39308 Human CDN
14	1091	100.0	1091	9	AdD38831 Human CDN
15	1091	100.0	1091	9	AdD40262 Human CDN
16	1091	100.0	1091	9	AdE50483 Human CDN
17	1091	100.0	1091	9	AdE20095 Human CDN
18	1091	100.0	1091	9	AdE50006 Human CDN
19	1091	100.0	1091	9	AdE21564 Human CDN
20	968	88.7	1185	4	Aah98653 Human EST
21	746.6	68.4	747	7	Aal59914 Human kal
22	727	66.6	945	6	AbK94900 Human nov
23	506.4	46.4	840	2	Aax80906 Human cdn

24	493.2	45.2	619	6	ABK30369
25	373.2	34.2	11820	3	AAA95944
26	297	27.2	405	5	AAF65654
27	219.4	20.1	1146	2	AAV84589
28	219.4	20.1	1146	4	ABAB3430
29	219.4	20.1	1146	8	ACH04931
30	219.4	20.1	1146	8	ACD44741
31	217.8	20.0	1166	4	AAD14841
32	217.6	19.9	1105	2	AAZ22638
33	217.6	19.9	1158	2	AAZ22639
34	217.6	19.9	1186	6	ABK92131
35	217.6	19.9	1186	7	ABX76468
36	217.6	19.9	1186	9	ADB80595
37	217.6	19.9	1186	9	ADB75387
38	217.6	19.9	1192	4	AAD14842
39	217.6	19.9	1204	3	AAA37072
40	217.6	19.9	1204	4	AAF54320
41	217.6	19.9	1204	4	AAZ21496
42	217.6	19.9	1204	6	ABK33628
43	217.6	19.9	1204	6	ABL88175
44	217.6	19.9	1204	6	ABL95664
45	217.6	19.9	1204	7	ACD24105

## ALIGNMENTS

## RESULT 1

AAA77671

ID AAA77671 standard; cDNA; 1091 BP.

XX

AC AAA77671;

XX

DT 07-NOV-2000 (first entry)

XX

DE Human PRO1303 cDNA sequence SEQ ID NO:202.

XX

KW Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;  
diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;  
KW angiogenic; proliferative; cardiant; cardiovascular; antiatherosclerotic;  
KW cytostatic; Gene therapy; vaccine; ss.

OS Homo sapiens.

XX

PN WO200032221-A2.

XX

PD 08-JUN-2000.

XX

PF 30-NOV-1999; 99WO-US028313.

XX

PR 01-DEC-1998; 98WO-US025108.

PR

PR 16-DEC-1998; 98US-0112850P.

PR

PR 12-JAN-1999; 99US-0115554P.

PR

PR 08-MAR-1999; 99WO-US005028.

PR

PR 12-MAR-1999; 99US-0123957P.

PR

PR 28-APR-1999; 99US-013445P.

PR

PR 14-MAY-1999; 99WO-US012252.

PR

PR 02-JUN-1999; 99US-0141037P.

PR

PR 20-JUL-1999; 99US-0144758P.

PR

PR 26-JUL-1999; 99US-0145698P.

PR

PR 01-SEP-1999; 99WO-US020111.

PR

PR 08-SEP-1999; 99WO-US020594.

PR

PR 13-SEP-1999; 99WO-US020944.

PR

PR 15-SEP-1999; 99WO-US021090.

PR

PR 15-SEP-1999; 99WO-US021547.

PR

PR 05-OCT-1999; 99WO-US023089.

PR

PR 29-OCT-1999; 99US-0162506P.

XX

(GETH ) GENENTECH INC.

PA

XX

PI Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ;

PI Goddard A, Godowski PU, Gurney AL, Klein RD, Kuo SS, Paoni NF;





PR 29-OCT-1998; 98US-0106384P.  
 PR 29-OCT-1998; 98US-0108500P.  
 PR 30-OCT-1998; 98US-0106464P.  
 PR 03-NOV-1998; 98US-0106856P.  
 PR 03-NOV-1998; 98US-0106902P.  
 PR 03-NOV-1998; 98US-0106905P.  
 PR 03-NOV-1998; 98US-0106919P.  
 PR 03-NOV-1998; 98US-0106932P.  
 PR 03-NOV-1998; 98US-0106934P.  
 PR 10-NOV-1998; 98US-0107783P.  
 PR 17-NOV-1998; 98US-0108775P.  
 PR 17-NOV-1998; 98US-0108779P.  
 PR 17-NOV-1998; 98US-0108787P.  
 PR 17-NOV-1998; 98US-0108788P.  
 PR 17-NOV-1998; 98US-0108801P.  
 PR 17-NOV-1998; 98US-0108802P.  
 PR 17-NOV-1998; 98US-0108806P.  
 PR 17-NOV-1998; 98US-0108807P.  
 PR 17-NOV-1998; 98US-0108867P.  
 PR 17-NOV-1998; 98US-0108925P.  
 PR 18-NOV-1998; 98US-0108848P.  
 PR 18-NOV-1998; 98US-0108849P.  
 PR 18-NOV-1998; 98US-0108850P.  
 PR 18-NOV-1998; 98US-0108851P.  
 PR 18-NOV-1998; 98US-0108852P.  
 PR 18-NOV-1998; 98US-0108858P.  
 PR 18-NOV-1998; 98US-0108904P.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood MI;  
 PI WPI; 2000-237871/20.  
 XX P-PSDB; AA199393.  
 DR  
 XX New mammalian DNA sequences encoding transmembrane, receptor or secreted  
 PT PRO polypeptides, useful for screening of potential peptide or small  
 PT molecule inhibitors of the relevant receptor/ligand interactions.  
 XX  
 PS Claim 2; Fig 107; 773pp; English.  
 XX  
 CC AAA37022 to AAA37144 encode the new isolated human transmembrane,  
 CC receptor or secreted PRO polypeptides given in AA199340 to AA199462. The  
 CC transmembrane and receptor PRO proteins can be used for screening of  
 CC potential peptide or small molecule inhibitors of the relevant  
 CC receptor/ligand interactions. The polypeptides and nucleotide sequences  
 CC encoding them have various industrial applications, including uses as  
 CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent PCR  
 CC primers and hybridisation probes used in the isolation of the PRO  
 CC polypeptides from the present invention  
 XX  
 SQ Sequence 1091 BP; 221 A; 335 C; 313 G; 222 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 1091; DB 3; Length 1091;  
 Best Local Similarity 100.0%; Pred.No. 1.5e-247;  
 Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAAGCAGGTCAATCCCTTGGTGACCTTCAGAGAAAGCAGAGAGGCGAGAGGTGGGGGC 60  
 Db 1 CAAGCAGGTCAATCCCTTGGTGACCTTCAGAGAAAGCAGAGAGGCGAGAGGTGGGGGC 60  
 QY 61 ACAGGGAAGGGTGACCTTCAGAGTTCCCTTTCCCCCAGACTTTGGAAGTAGCCACC 120  
 Db 61 ACAGGGAAGGGTGACCTTCAGAGTTCCCTTTCCCCCAGACTTTGGAAGTAGCCACC 120  
 QY 121 ATGGGGCTCAGCATCTTTTTTGTCTTCTTGGCTCAGCAGGCGACACCG 180  
 Db 121 ATGGGGCTCAGCATCTTTTTTGTCTTCTTGGCTCAGCAGGCGACACCG 180  
 QY 181 AAGATTTCATGGCACTAGTGCGGCGTAACCTCACAGCCGTGGCAGGTGGGGTGT 240  
 Db 181 AAGATTTCATGGCACTAGTGCGGCGTAACCTCACAGCCGTGGCAGGTGGGGTGT 240

241	QY	GAGGGCACAGCGCTGCGCTGCGGGGTGTCCTTATTGACACACAGGTGGGTCTCTACACGG	300
241	Db	GAGGGCACAGCGCTGCGGTCTGCGGGGTGTCCTTATTGACACACAGGTGGGTCTCTACACGG	300
301	QY	GCTCACTGACGGCAGCAGGTACTGGGTGCGCTGGGGAAACACAGCCTCAGCAGTCC	360
301	Db	GCTCACTGACGGCAGCAGGTACTGGGTGCGCTGGGGAAACACAGCCTCAGCAGTCC	360
361	QY	GACTGACCGAGCAGATCCGGCACAGCGGTTCTCTGTGACCCATCCGGTACTCTGGGA	420
361	Db	GACTGACCGAGCAGATCCGGCACAGCGGTTCTCTGTGACCCATCCGGTACTCTGGGA	420
421	QY	GCTTCGACGAGCAGACGACCTCCGGCTGTGCGGCTGCGCTGCGCTGCGCTGCGCGTA	480
421	Db	GCTTCGACGAGCAGACGACCTCCGGCTGTGCGGCTGCGCTGCGCTGCGCTGCGCGTA	480
481	QY	ACGACGAGCGTTCAACCCCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG	540
481	Db	ACGACGAGCGTTCAACCCCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG	540
541	QY	CACGTCTCAGGCTGGGGCATCAACCAACACCGAACCCATTCCCGATCTGCTCCAG	600
541	Db	CACGTCTCAGGCTGGGGCATCAACCAACACCGAACCCATTCCCGATCTGCTCCAG	600
601	QY	TGCTCAACCTTCCATCGTCTCCCATGCCACTGCGATGTGTGTATCCCGGGAGAATC	660
601	Db	TGCTCAACCTTCCATCGTCTCCCATGCCACTGCGATGTGTGTATCCCGGGAGAATC	660
661	QY	ACGACCAACATGGTGTGTGTCAGGCGGCGTCCCGGGCAGGATGCTGCGAGGTGATTC	720
661	Db	ACGACCAACATGGTGTGTGTCAGGCGGCGTCCCGGGCAGGATGCTGCGAGGTGATTC	720
721	QY	GGGGGCCCCCTGGTGTGTGGGGAGTCTCTCAAGGTCTGCTGCTGGGGTCTGTGGG	780
721	Db	GGGGGCCCCCTGGTGTGTGGGGAGTCTCTCAAGGTCTGCTGCTGGGGTCTGTGGG	780
781	QY	CCCTGTGGAACAGATGGCATCCCTGGAGTCTACACCTATATTGCAAGTATGTGGACTG	840
781	Db	CCCTGTGGAACAGATGGCATCCCTGGAGTCTACACCTATATTGCAAGTATGTGGACTG	840
841	QY	ATCCGATGATCATGAGGAACAACCTGACCTGTTTCTCCAGCTCCACCCCAACCCCTTAA	900
841	Db	ATCCGATGATCATGAGGAACAACCTGACCTGTTTCTCCAGCTCCACCCCAACCCCTTAA	900
901	QY	CTTTGGTAACCCCTCTGGCCCTCAGACACCAATATCTCTCATCTCCCTAGCTCC	960
901	Db	CTTTGGTAACCCCTCTGGCCCTCAGACACCAATATCTCTCATCTCCCTAGCTCC	960
961	QY	ACTCTTTGTCGCTGGGAACCTTCTTGAACCTTTAACTCCTGCGACCCCTCTTAAGACCCA	1020
961	Db	ACTCTTTGTCGCTGGGAACCTTCTTGAACCTTTAACTCCTGCGACCCCTCTTAAGACCCA	1020
1021	QY	CGAGCGGGTGAGAGAAGTGTGCAATAGTCTGGAAATAAATAAATGAAGGGGGGCAAA	1080
1021	Db	CGAGCGGGTGAGAGAAGTGTGCAATAGTCTGGAAATAAATAAATGAAGGGGGGCAAA	1080
1081	QY	AAAAAAAAAAAA 1091	
1081	Db	AAAAAAAAAAAA 1091	

RESULT 4  
AAF54341  
ID AAF54341 standard; DNA; 1091 BP.  
XX  
AC AAF54341;  
XX  
AC  
XX  
DT 02-APR-2001 (first entry)  
XX  
DE DNA encoding protein of the invention #54.  
XX  
XX  
KW Secreted; transmembrane; gene therapy; ss.  
XX

Unidentified.	
WO200078961-A1.	
28-DEC-2000.	
18-FEB-2000; 2000WO-US004342.	
23-JUN-1999; 99US-0141037P.	
20-JUL-1999; 99US-0144758P.	
26-JUL-1999; 99US-0145698P.	
01-SEP-1999; 99WO-US020111.	
29-OCT-1999; 99US-0162506P.	
30-NOV-1999; 99WO-US028313.	
02-DEC-1999; 99WO-US028551.	
16-DEC-1999; 99WO-US030095.	
05-JAN-2000; 2000WO-US000219.	
06-JAN-2000; 2000WO-US000376.	
(GETH ) GENENTECH INC.	
Baker KP, Botstein D, Deenoyers L, Eaton DL, Ferrara N, Fong S;	
Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AU, Hillan KJ;	
Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;	
Williams PM, Wood WI;	
WPI; 2001-071395/08.	
Secreted and transmembrane proteins and nucleic acids designated PRO,	
useful as hybridization probes, in chromosome and gene mapping and gene	
therapy.	
Claim 2; Fig 107; 787pp; English.	
The present invention relates to secreted and transmembrane proteins.	
These proteins and the DNA encoding them may be used as hybridization	
probes, in chromosome and gene mapping and in the generation of anti-	
sense RNA and DNA. They may also be used to generate either	
transgenic animals or knockout animals which are in turn useful for	
development and screening of therapeutically useful reagents. The nucleic	
acids may also be used in gene therapy	

XX	Sequence	1091 BP;	221 A;	335 C;	313 G;	222 T;	0 U;	0 Other;		
SQL	Query Match	100.0%;	Score	1091;	DB	4;	Length	1091;		
	Best Local Similarity	100.0%;	Pred	No.	1.5e-247;					
	Matches 1091;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	
QY	1	CAAGCAGGT	CATCCCTTGGTG	ACCTTCA	AAGAGAG	CAGAGG	CAGAGG	CAGAGG	TGGGGGGC	60
Db	1	CAAGCAGGT	CATCCCTTGGTG	ACCTTCA	AAGAGAG	CAGAGG	CAGAGG	CAGAGG	TGGGGGGC	60
QY	61	ACAGGGA	AGGTCGACTCTG	AGATTCC	CTTTCC	CCAGAC	ATTTGG	AAGTAC	CCAC	120
Db	61	ACAGGGA	AGGTCGACTCTG	AGATTCC	CTTTCC	CCAGAC	ATTTGG	AAGTAC	CCAC	120
QY	121	ATGGGGCT	CAGCAATCTTTT	TCTCTCTG	TGTCTT	TGGGCT	CAGCC	AGGAG	CCAC	180
Db	121	ATGGGGCT	CAGCAATCTTTT	TCTCTCTG	TGTCTT	TGGGCT	CAGCC	AGGAG	CCAC	180
QY	181	AGATTTT	CAATGG	CAC	TGAGT	TGGGCG	TAACT	CAC	AGCCG	240
Db	181	AGATTTT	CAATGG	CAC	TGAGT	TGGGCG	TAACT	CAC	AGCCG	240
QY	241	GAGGGC	AC	AGCCTCG	CTGGGG	GGTTC	CTTATT	GAC	CAC	300
Db	241	GAGGGC	AC	AGCCTCG	CTGGGG	GGTTC	CTTATT	GAC	CAC	300
QY	301	GCTCA	CTG	CAG	CGG	CAG	CAG	GTACT	TGGG	360
Db	301	GCTCA	CTG	CAG	CGG	CAG	CAG	GTACT	TGGG	360
QY	361	GACTG	AC	CCG	AG	CAG	ATCC	GCCTCT	CTGT	420
Db	361	GACTG	AC	CCG	AG	CAG	ATCC	GCCTCT	CTGT	420

XX	OS	Homo sapiens.
XX	XX	US2003073130-A1.
XX	XX	17-APR-2003.
XX	XX	11-DEC-2001; 2001US-00015869.
XX	XX	01-SEP-1998; 98US-0098716P;
XX	XX	01-SEP-1998; 98US-0098723P;
XX	XX	01-SEP-1998; 98US-0098749P;
XX	XX	01-SEP-1998; 98US-0098750P;
XX	XX	02-SEP-1998; 98US-0098803P;
XX	XX	02-SEP-1998; 98US-0098821P;
XX	XX	02-SEP-1998; 98US-0098843P;
XX	XX	02-SEP-1998; 98US-0098936P;
XX	XX	09-SEP-1998; 98US-0099598P;
XX	XX	09-SEP-1998; 98US-0099602P;
XX	XX	09-SEP-1998; 98US-0099642P;
XX	XX	10-SEP-1998; 98US-0099741P;
XX	XX	10-SEP-1998; 98US-0099754P;
XX	XX	10-SEP-1998; 98US-0099763P;
XX	XX	10-SEP-1998; 98US-0099792P;
XX	XX	10-SEP-1998; 98US-0099812P;
XX	XX	10-SEP-1998; 98US-0099815P;
XX	XX	10-SEP-1998; 98US-0099816P;
XX	XX	10-SEP-1998; 98US-0100385P;
XX	XX	15-SEP-1998; 98US-0100388P;
XX	XX	15-SEP-1998; 98US-0100390P;
XX	XX	16-SEP-1998; 98US-0100584P;
XX	XX	16-SEP-1998; 98US-0100627P;
XX	XX	16-SEP-1998; 98US-0100661P;
XX	XX	16-SEP-1998; 98US-0100662P;
XX	XX	16-SEP-1998; 98US-0100664P;
XX	XX	17-SEP-1998; 98US-0100683P;
XX	XX	17-SEP-1998; 98US-0100684P;
XX	XX	17-SEP-1998; 98US-0100710P;
XX	XX	17-SEP-1998; 98US-0100711P;
XX	XX	17-SEP-1998; 98US-0100930P;
XX	XX	17-SEP-1998; 98US-0100930P;
XX	XX	18-SEP-1998; 98US-0100848P;
XX	XX	18-SEP-1998; 98US-0100849P;
XX	XX	18-SEP-1998; 98US-0101014P;
XX	XX	18-SEP-1998; 98US-0101068P;
XX	XX	18-SEP-1998; 98US-0101071P;
XX	XX	22-SEP-1998; 98US-0101279P;
XX	XX	23-SEP-1998; 98US-0101471P;
XX	XX	23-SEP-1998; 98US-0101472P;
XX	XX	23-SEP-1998; 98US-0101474P;
XX	XX	23-SEP-1998; 98US-0101475P;
XX	XX	23-SEP-1998; 98US-0101476P;
XX	XX	23-SEP-1998; 98US-0101477P;
XX	XX	23-SEP-1998; 98US-0101479P;
XX	XX	24-SEP-1998; 98US-0101738P;
XX	XX	24-SEP-1998; 98US-0101741P;
XX	XX	24-SEP-1998; 98US-0101915P;
XX	XX	24-SEP-1998; 98US-0101916P;
XX	XX	29-SEP-1998; 98US-0102207P;
XX	XX	29-SEP-1998; 98US-0102240P;
XX	XX	29-SEP-1998; 98US-0102307P;
XX	XX	29-SEP-1998; 98US-0102330P;
XX	XX	29-SEP-1998; 98US-0102331P;
XX	XX	30-SEP-1998; 98US-0102484P;
XX	XX	30-SEP-1998; 98US-0102487P;
XX	XX	30-SEP-1998; 98US-0102570P;
XX	XX	30-SEP-1998; 98US-0102571P;
XX	XX	01-OCT-1998; 98US-0102684P;
XX	XX	01-OCT-1998; 98US-0102687P;
XX	XX	02-OCT-1998; 98US-0102965P;
XX	XX	06-OCT-1998; 98US-0103258P;

PR	06-OCT-1998;	98US-0103449P.	PR	11-FEB-2000;	2000WO-US003565.
PR	07-OCT-1998;	98US-0103314P.	PR	18-FEB-2000;	2000WO-US004342.
PR	07-OCT-1998;	98US-0103315P.	PR	24-FEB-2000;	2000WO-US005004.
PR	07-OCT-1998;	98US-0103328P.	PR	02-MAR-2000;	2000WO-US005941.
PR	07-OCT-1998;	98US-0103395P.	PR	15-MAR-2000;	2000WO-US006884.
PR	07-OCT-1998;	98US-0103396P.	PR	17-MAY-2000;	2000WO-US013705.
PR	07-OCT-1998;	98US-0103401P.	PR	22-MAY-2000;	2000WO-US014042.
PR	08-OCT-1998;	98US-0103633P.	PR	30-MAY-2000;	2000WO-US014941.
PR	08-OCT-1998;	98US-0103678P.	PR	23-JUN-2000;	2000WO-US015264.
PR	08-OCT-1998;	98US-0103679P.	PR	23-AUG-2000;	2000WO-US023522.
PR	08-OCT-1998;	98US-0103711P.	PR	24-AUG-2000;	2000WO-US023328.
PR	14-OCT-1998;	98US-0104257P.	PR	08-NOV-2000;	2000WO-US030952.
PR	20-OCT-1998;	98US-0105000P.	PR	10-NOV-2000;	2000WO-US030873.
PR	20-OCT-1998;	98US-0105002P.	PR	01-DEC-2000;	2000WO-US032678.
PR	21-OCT-1998;	98US-0105024P.	PR	28-FEB-2001;	2001WO-US006520.
PR	22-OCT-1998;	98US-0105169P.	PR	01-MAR-2001;	2001WO-US006666.
PR	22-OCT-1998;	98US-0105266P.	PR	01-JUN-2001;	2001WO-US017800.
PR	26-OCT-1998;	98US-0105693P.	PR	20-JUN-2001;	2001WO-US019692.
PR	26-OCT-1998;	98US-0105694P.	PR	29-JUN-2001;	2001WO-US021066.
PR	27-OCT-1998;	98US-0105807P.	PR	09-JUL-2001;	2001WO-US021735.
PR	27-OCT-1998;	98US-0105881P.	PR	04-SEP-2001;	2001US-00946374.
PR	27-OCT-1998;	98US-0105882P.	XX		
PR	27-OCT-1998;	98US-0106062P.	XX		
PR	28-OCT-1998;	98US-0106023P.	PI	Baker KP, Botstein D, Deanovers L, Eaton JC, Ferrara N, Fong S;	
PR	28-OCT-1998;	98US-0106029P.	PI	Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;	
PR	28-OCT-1998;	98US-0106030P.	PI	Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;	
PR	28-OCT-1998;	98US-0106032P.	PI	Williams PW, Wood WI;	
PR	28-OCT-1998;	98US-0106033P.	XX		
PR	28-OCT-1998;	98US-0106178P.	DR	WPI: 2003-585293/55.	
PR	29-OCT-1998;	98US-0106248P.	DR	P-PSDB; AB033635.	
PR	29-OCT-1998;	98US-0106384P.	XX		
PR	29-OCT-1998;	98US-0108500P.	PT	Novel isolated PRO polypeptides e.g. PRO1130, PRO1275, PRO1418, PRO1555,	
PR	30-OCT-1998;	98US-0108646P.	PT	PRO1787 that modulate glucose or free fatty acid uptake by skeletal	
PR	03-NOV-1998;	98US-0108566P.	PT	muscle cells, and are useful for treating diabetes, hyper- or hypo-	
PR	03-NOV-1998;	98US-0106302P.			
PR	03-NOV-1998;	98US-0106919P.			
PR	03-NOV-1998;	98US-0106932P.			
PR	03-NOV-1998;	98US-0106934P.			
PR	10-NOV-1998;	98US-0107783P.			
PR	17-NOV-1998;	98US-0108775P.			
PR	17-NOV-1998;	98US-0108779P.			
PR	17-NOV-1998;	98US-0108787P.			
PR	17-NOV-1998;	98US-0108788P.			
PR	17-NOV-1998;	98US-0108801P.			
PR	17-NOV-1998;	98US-0108802P.			
PR	17-NOV-1998;	98US-0108806P.			
PR	17-NOV-1998;	98US-0108807P.			
PR	17-NOV-1998;	98US-0108867P.			
PR	17-NOV-1998;	98US-0108925P.			
PR	18-NOV-1998;	98US-0108848P.			
PR	18-NOV-1998;	98US-0108849P.			
PR	18-NOV-1998;	98US-0108850P.			
PR	18-NOV-1998;	98US-0108851P.			
PR	18-NOV-1998;	98US-0108852P.			
PR	18-NOV-1998;	98US-0108858P.			
PR	18-NOV-1998;	98US-0108904P.			
PR	22-DEC-1998;	98US-0113296P.			
PR	05-JAN-1999;	98US-0114223P.			
PR	16-APR-1999;	98US-0129674P.			
PR	23-JUN-1999;	98US-0141037P.			
PR	20-JUL-1999;	98US-01447			



[illegible]

PR	08-OCT-1998;	98US-0103679P.	PR	23-AUG-2000;	2000WO-US023522.
PR	08-OCT-1998;	98US-0103711P.	PR	24-AUG-2000;	2000WO-US023328.
PR	14-OCT-1998;	98US-0104257P.	PR	08-NOV-2000;	2000WO-US030952.
PR	20-OCT-1998;	98US-0104987P.	PR	10-NOV-2000;	2000WO-US030873.
PR	20-OCT-1998;	98US-0105000P.	PR	01-DEC-2000;	2000WO-US032678.
PR	20-OCT-1998;	98US-0105002P.	PR	28-FEB-2001;	2001WO-US006520.
PR	21-OCT-1998;	98US-0105104P.	PR	01-MAR-2001;	2001WO-US006666.
PR	22-OCT-1998;	98US-0105169P.	PR	01-JUN-2001;	2001WO-US017800.
PR	22-OCT-1998;	98US-0105266P.	PR	20-JUN-2001;	2001WO-US019692.
PR	26-OCT-1998;	98US-0105693P.	PR	29-JUN-2001;	2001WO-US021066.
PR	26-OCT-1998;	98US-0105694P.	PR	09-JUL-2001;	2001WO-US021735.
PR	27-OCT-1998;	98US-0105807P.	PR	04-SEP-2001;	2001US-00946374.
PR	27-OCT-1998;	98US-0105881P.	XX	(GETH )	GENENTECH INC.
PR	27-OCT-1998;	98US-0105882P.	XX		
PR	27-OCT-1998;	98US-0106062P.	XX		
PR	28-OCT-1998;	98US-0106023P.	PI	Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;	
PR	28-OCT-1998;	98US-0106029P.	PI	Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;	
PR	28-OCT-1998;	98US-0106030P.	PI	Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;	
PR	28-OCT-1998;	98US-0106032P.	PI	Williams PM, Wood WT;	
PR	28-OCT-1998;	98US-0106033P.	XX		
PR	28-OCT-1998;	98US-0106178P.	DR	WPI; 2003-492259/46.	
PR	29-OCT-1998;	98US-0106248P.	DR	P-PSDB; ABO44488.	
PR	29-OCT-1998;	98US-0106384P.	XX		
PR	29-OCT-1998;	98US-0108500P.	XX		
PR	30-OCT-1998;	98US-0106464P.	PT	Novel secreted and transmembrane polypeptides and polynucleotides	
PR	30-OCT-1998;	98US-0106856P.	PT	encoding them useful for treating various cardiac insufficiency	
PR	03-NOV-1998;	98US-0106902P.	PT	disorders, bone and/or cartilage disorders such as sports injuries and	
PR	03-NOV-1998;	98US-0106905P.	PT	arthritis.	
PR	03-NOV-1998;	98US-0106919P.			
PR	03-NOV-1998;	98US-0106932P.			
PR	03-NOV-1998;	98US-0106934P.			
PR	10-NOV-1998;	98US-0107783P.			
PR	17-NOV-1998;	98US-0108775P.			
PR	17-NOV-1998;	98US-0108779P.			
PR	17-NOV-1998;	98US-0108787P.			
PR	17-NOV-1998;	98US-0108801P.			
PR	17-NOV-1998;	98US-0108802P.			
PR	17-NOV-1998;	98US-0108806P.			
PR	17-NOV-1998;	98US-0108807P.			
PR	17-NOV-1998;	98US-0108867P.			
PR	17-NOV-1998;	98US-0108925P.			
PR	18-NOV-1998;	98US-0108848P.			
PR	18-NOV-1998;	98US-0108849P.			
PR	18-NOV-1998;	98US-0108850P.			
PR	18-NOV-1998;	98US-0108851P.			
PR	18-NOV-1998;	98US-0108852P.			
PR	18-NOV-1998;	98US-0108859P.			
PR	18-NOV-1998;	98US-0108904P.			
PR	22-DEC-1998;	98US-0113296P.			
PR	30-DEC-1998;	98US-0114223P.			
PR	05-JAN-1999;	98US-0108510P.			
PR	16-APR-1999;	98US-0129674P.			
PR	23-JUN-1999;	98US-0141037P.			
PR	26-JUL-1999;	98US-0144758P.			
PR	01-SEP-1999;	98US-0145698P.			
PR	15-SEP-1999;	98US-0202111P.			
PR	29-OCT-1999;	98US-0202119P.			
PR	30-NOV-1999;	98US-0162506P.			
PR	02-DEC-1999;	98US-02028313.			
PR	16-DEC-1999;	98US-02028551.			
PR	05-JAN-2000;	98US-03030095.			
PR	11-FEB-2000;	2000WO-US000219.			
PR	24-FEB-2000;	2000WO-US000376.			
PR	02-MAR-2000;	2000WO-US000365.			
PR	15-MAR-2000;	2000WO-US000342.			
PR	17-MAY-2000;	2000WO-US000504.			
PR	30-MAY-2000;	2000WO-US0005841.			
PR	02-J				



XX	(GETH ) GENENTECH INC.
PA	Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
XX	PI Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
PI	Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
PI	Williams PM, Wood WI;
XX	WPI; 2003-585292/55.
DR	P-PSDB; ABO33512.
XX	
PT	Novel isolated PRO polypeptides e.g. PRO1491 and PRO1571, useful in the
PT	preparation of a medicament for treating a condition responsive to PRO
PT	polypeptide, and as therapeutic agents e.g. vaccines.
XX	
PS	Claim 2; Fig 107; 561pp; English.
XX	
CC	The invention describes an isolated PRO (secreted and transmembrane)
CC	polypeptide (I), having at least 80% sequence identity to a sequence
CC	
	Query Match      100.0%; Score 1091; DB 8; Length 1091;
	Best Local Similarity 100.0%; Pred. No. 1.5e-247; Indels 0; Gaps 0;
	Matches 1091; Conservative 0; Mismatches 0;
QY	1 CAAGCAGGTCATCCCCTTGCTGACCTTCAAAGAGAAGCAGAGGGCAGAGGTGGGGGC 60
Db	1 CAAGCAGGTCATCCCCTTGCTGACCTTCAAGAGAAGCAGAGGGCAGAGGTGGGGGC 60
QY	61 ACAGGAAAGGGGTGACTCTGAGATTCCCTTTTCCCCCAGACTTTGGAAAGTACCACCC 120
Db	61 ACAGGAAAGGGGTGACTCTGAGATTCCCTTTTCCCCCAGACTTTGGAAAGTACCACCC 120
QY	121 ATGGGGCTCAGCATCTTTTGTCTCTGTGTCTTGGGCTCAGCCAGGCGAGCACACCG 180
Db	121 ATGGGGCTCAGCATCTTTTGTCTCTGTGTCTTGGGCTCAGCCAGGCGAGCACACCG 180
QY	181 AAGATTTTCAATGGCACTGAGTGTGGGCGTAACCTCACAGCCGTGGCAGGTGGGGCTGTTT 240
Db	181 AAGATTTTCAATGGCACTGAGTGTGGGCGTAACCTCACAGCCGTGGCAGGTGGGGCTGTTT 240
QY	241 GAGGCGACCAAGCCTGGCTCGGGGGTGTCTTTATTGACACAGGTGGGTCTCTCACAGCG 300
Db	241 GAGGCGACCAAGCCTGGCTCGGGGGTGTCTTTATTGACACAGGTGGGTCTCTCACAGCG 300
QY	301 GCTCACTGACGGCAGCAGGTACTGGGTGGCTGGGGGAACAACACCCCTCAGCCAGCTC 360
Db	301 GCTCACTGACGGCAGCAGGTACTGGGTGGCTGGGGGAACAACACCCCTCAGCCAGCTC 360
QY	361 GACTGGACCGAGCAGATCCGGCA CAGCGGCTTCTCTGTGACCCAATCCCGGTACTCTGGGA 420
Db	361 GACTGGACCGAGCAGATCCGGCA CAGCGGCTTCTCTGTGACCCAATCCCGGTACTCTGGGA 420
QY	421 GCCTCGAAGCCACGAGCAGCA CTCTCGGGTGTCTGGGCTGCCGCTGCCGTCGCGGTA 480
Db	421 GCCTCGAAGCCACGAGCAGCA CTCTCGGGTGTCTGGGCTGCCGCTGCCGTCGCGGTA 480
QY	481 ACCAGCAGCGTTCAACCCCTGCCCTGCCCAATGACTGTGCAA CGGTGGCACCGAGTGC 540
Db	481 ACCAGCAGCGTTCAACCCCTGCCCTGCCCAATGACTGTGCAA CGGTGGCACCGAGTGC 540
QY	541 CACGTCTCAGGCTGGGGCATCACCAACCCACCGAACCATTGCCCGGATGTGCTCCAG 600
Db	541 CACGTCTCAGGCTGGGGCATCACCAACCCACCGAACCATTGCCCGGATGTGCTCCAG 600
QY	601 TGCCCTCAACCTCTCCATCGTCTCCCATGCCACTGCCATGTGTGTATCCCGGGAGAATC 660
Db	601 TGCCCTCAACCTCTCCATCGTCTCCCATGCCACTGCCATGTGTGTATCCCGGGAGAATC 660
QY	661 ACGAGCAACATGGTGTGTGACGGGGCGTCCCGGGCAGGATGCCCTGCCAGGGTGAATCT 720
Db	661 ACGAGCAACATGGTGTGTGACGGGGCGTCCCGGGCAGGATGCCCTGCCAGGGTGAATCT 720
QY	721 GGGGGCCCCCTGCTGTGTGGGGGAGTCCCTTCAAGGTCGTGGNCTCTGGGGGTCTGTGGGG 780

Mon Jun 21 09:00:25 2004

us-10-015-385a-193.rng

```
Db      721  |||||GGGGCCCCCTGGTGTGGGGAGTCCCTTAAGTCTGGTGTCTGGGGTCTGTGGG 780
Qy      781  CCCTGTGGACAAGATGCGATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGGACTGG 840
Db      781  CCCTGTGGACAAGATGCGATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGGACTGG 840
Qy      841  ATCCGATGATCATGAGGACCACTGACCTGTTTCTCCACCTCCACCCCAACCCCTTAA 900
Db      841  ATCCGATGATCATGAGGACCACTGACCTGTTTCTCCACCTCCACCCCAACCCCTTAA 900
Qy      901  CTTGGGTACCCCTCTGCGCTCGAGCACCAATATCTCTCCATCACTTCCCTAGTCTCC 960
Db      901  CTTGGGTACCCCTCTGCGCTCGAGCACCAATATCTCTCCATCACTTCCCTAGTCTCC 960
Qy      961  ACTCTTGTGGCTGGGAACCTCTTGGAACTTTTGAACCTTTAACTCTGCGAGCCCTTCTAAGACCCA 1020
Db      961  ACTCTTGTGGCTGGGAACCTCTTGGAACTTTTGAACCTTTAACTCTGCGAGCCCTTCTAAGACCCA 1020
Qy      1021  CGAGCGGCTGAGAGAGTGTGCAATAGTCTGGAATAAATAATAAATGAAGAGGGGCAAA 1080
Db      1021  CGAGCGGCTGAGAGAGTGTGCAATAGTCTGGAATAAATAATAAATGAAGAGGGGCAAA 1080
Qy      1081  AAAAAAAAAA 1091
Db      1081  AAAAAAAAAA 1091

RESULT 8
ADCI18062
ID  ADCI18062 standard, cDNA; 1091 BP.
AC  ADCI18062;
XX
XX  18-DEC-2003 (first entry)
DE  Human PRO polynucleotide #54.
KW  Human; PRO; gene; ss; protein electrophoresis; chromosome mapping;
KW  gene mapping; genetic disorder.
XX
XX  Homo sapiens.
XX
XX  US2003064925-A1.
XX
XX  03-APR-2003.
XX
XX  10-DEC-2001; 2001US-00013907.
XX
XX  01-SEP-1998; 98US-0098716P.
XX  01-SEP-1998; 98US-0098723P.
XX  01-SEP-1998; 98US-0098749P.
XX  01-SEP-1998; 98US-0098750P.
XX  02-SEP-1998; 98US-0098803P.
XX  02-SEP-1998; 98US-0098821P.
XX  02-SEP-1998; 98US-0098843P.
XX  09-SEP-1998; 98US-0099536P.
XX  09-SEP-1998; 98US-0099596P.
XX  09-SEP-1998; 98US-0099598P.
XX  09-SEP-1998; 98US-0099602P.
XX  09-SEP-1998; 98US-0099642P.
XX  10-SEP-1998; 98US-0099741P.
XX  10-SEP-1998; 98US-0099754P.
XX  10-SEP-1998; 98US-0099763P.
XX  10-SEP-1998; 98US-0099792P.
XX  10-SEP-1998; 98US-0099808P.
XX  10-SEP-1998; 98US-0099812P.
XX  10-SEP-1998; 98US-0099815P.
XX  10-SEP-1998; 98US-0099816P.
XX  15-SEP-1998; 98US-0100385P.
XX  15-SEP-1998; 98US-0100389P.
XX  15-SEP-1998; 98US-0100390P.
XX  16-SEP-1998; 98US-0100584P.
XX  16-SEP-1998; 98US-0100627P.
XX  16-SEP-1998; 98US-0100661P.
XX  16-SEP-1998; 98US-0100662P.
XX  16-SEP-1998; 98US-0100664P.
XX  17-SEP-1998; 98US-0100683P.
XX  17-SEP-1998; 98US-0100684P.
XX  17-SEP-1998; 98US-0100710P.
XX  17-SEP-1998; 98US-0100711P.
XX  17-SEP-1998; 98US-0100919P.
XX  17-SEP-1998; 98US-0100930P.
XX  17-SEP-1998; 98US-0100948P.
XX  18-SEP-1998; 98US-0100849P.
XX  18-SEP-1998; 98US-0101014P.
XX  18-SEP-1998; 98US-0101068P.
XX  18-SEP-1998; 98US-0101071P.
XX  22-SEP-1998; 98US-0101279P.
XX  23-SEP-1998; 98US-0101471P.
XX  23-SEP-1998; 98US-0101472P.
XX  23-SEP-1998; 98US-0101474P.
XX  23-SEP-1998; 98US-0101476P.
XX  23-SEP-1998; 98US-0101477P.
XX  24-SEP-1998; 98US-0101479P.
XX  24-SEP-1998; 98US-0101738P.
XX  24-SEP-1998; 98US-0101741P.
XX  24-SEP-1998; 98US-0101743P.
XX  24-SEP-1998; 98US-0101915P.
XX  24-SEP-1998; 98US-0101916P.
XX  24-SEP-1998; 98US-010207P.
XX  29-SEP-1998; 98US-0102240P.
XX  29-SEP-1998; 98US-0102307P.
XX  29-SEP-1998; 98US-0102330P.
XX  29-SEP-1998; 98US-0102331P.
XX  30-SEP-1998; 98US-0102484P.
XX  30-SEP-1998; 98US-0102487P.
XX  30-SEP-1998; 98US-0102570P.
XX  30-SEP-1998; 98US-0102571P.
XX  01-OCT-1998; 98US-0102684P.
XX  01-OCT-1998; 98US-0102687P.
XX  02-OCT-1998; 98US-0102965P.
XX  06-OCT-1998; 98US-0103258P.
XX  06-OCT-1998; 98US-0103449P.
XX  07-OCT-1998; 98US-0103314P.
XX  07-OCT-1998; 98US-0103315P.
XX  07-OCT-1998; 98US-0103328P.
XX  07-OCT-1998; 98US-0103395P.
XX  07-OCT-1998; 98US-0103396P.
XX  07-OCT-1998; 98US-0103401P.
XX  08-OCT-1998; 98US-0103633P.
XX  08-OCT-1998; 98US-0103678P.
XX  08-OCT-1998; 98US-0103679P.
XX  14-OCT-1998; 98US-0104257P.
XX  20-OCT-1998; 98US-0104987P.
XX  20-OCT-1998; 98US-0105000P.
XX  20-OCT-1998; 98US-0105002P.
XX  21-OCT-1998; 98US-0105104P.
XX  22-OCT-1998; 98US-0105169P.
XX  22-OCT-1998; 98US-0105266P.
XX  26-OCT-1998; 98US-0105693P.
XX  26-OCT-1998; 98US-0105807P.
XX  27-OCT-1998; 98US-0105881P.
XX  27-OCT-1998; 98US-0105882P.
XX  27-OCT-1998; 98US-0106062P.
XX  28-OCT-1998; 98US-0106023P.
XX  28-OCT-1998; 98US-0106030P.
XX  28-OCT-1998; 98US-0106032P.
XX  28-OCT-1998; 98US-0106033P.
XX  28-OCT-1998; 98US-0106178P.
XX  29-OCT-1998; 98US-0106248P.
XX  29-OCT-1998; 98US-0106384P.
```

PR	29-OCT-1998;	98US-0108500P.	PT	Novel isolated PRO polypeptides e.g. PRO1491 and PRO1571, useful in the
PR	30-OCT-1998;	98US-0106464P.	PT	preparation of a medicament for treating a condition responsive to PRO
PR	03-NOV-1998;	98US-0106856P.	XX	polypeptide, and as therapeutic agents e.g. vaccines.
PR	03-NOV-1998;	98US-0106902P.	PS	Claim 2; SEQ ID NO 193; 555pp; English.
PR	03-NOV-1998;	98US-0106919P.	XX	
PR	03-NOV-1998;	98US-0106932P.	CC	The invention relates to human PRO polypeptides and the polynucleotides
PR	03-NOV-1998;	98US-0106934P.	CC	encoding them. The sequences are useful in the preparation of a
PR	10-NOV-1998;	98US-0107783P.	CC	medicament for treating a condition responsive to a PRO polypeptide. The
PR	10-NOV-1998;	98US-0108775P.	CC	polypeptides are useful in a number of functional biological assays, as
PR	17-NOV-1998;	98US-0108779P.	CC	molecular weight markers for protein electrophoresis and as therapeutic
PR	17-NOV-1998;	98US-0108787P.		
PR	17-NOV-1998;	98US-0108788P.		
PR	17-NOV-1998;	98US-0108801P.		
PR	17-NOV-1998;	98US-0108802P.		
PR	17-NOV-1998;	98US-0108806P.		
PR	17-NOV-1998;	98US-0108807P.		
PR	17-NOV-1998;	98US-0108867P.		
PR	17-NOV-1998;	98US-0108925P.		
PR	18-NOV-1998;	98US-0108848P.		
PR	18-NOV-1998;	98US-0108849P.		
PR	18-NOV-1998;	98US-0108850P.		
PR	18-NOV-1998;	98US-0108851P.		
PR	18-NOV-1998;	98US-0108852P.		
PR	18-NOV-1998;	98US-0108858P.		
PR	18-NOV-1998;	98US-0108904P.		
PR	22-DEC-1998;	98US-01113296P.		
PR	30-DEC-1998;	98US-0114223P.		
PR	05-JAN-1999;	98US-0114223P.		
PR	16-APR-1999;	98US-0129674P.		
PR	23-JUN-1999;	98US-0141037P.		
PR	20-JUL-1999;	98US-0144758P.		
PR	26-JUL-1999;	98US-0145698P.		
PR	01-SEP-1999;	98US-0145698P.		
PR	15-SEP-1999;	98US-0145698P.		
PR	29-OCT-1999;	98US-0162506P.		
PR	30-NOV-1999;	98US-0162506P.		
PR	02-DEC-1999;	98US-0162506P.		
PR	16-DEC-1999;	98US-0162506P.		
PR	05-JAN-2000;	2000WO-US000219.		
PR	06-JAN-2000;	2000WO-US000376.		
PR	11-FEB-2000;	2000WO-US003565.		
PR	18-FEB-2000;	2000WO-US004342.		
PR	24-FEB-2000;	2000WO-US005004.		
PR	02-MAR-2000;	2000WO-US005841.		
PR	15-MAR-2000;	2000WO-US006884.		
PR	22-MAR-2000;	2000WO-US013705.		
PR	17-MAY-2000;	2000WO-US014042.		
PR	30-MAY-2000;	2000WO-US014941.		
PR	02-JUN-2000;	2000WO-US015264.		
PR	23-AUG-2000;	2000WO-US023522.		
PR	24-AUG-2000;	2000WO-US023328.		
PR	08-NOV-2000;	2000WO-US030952.		
PR	10-NOV-2000;	2000WO-US030873.		
PR	01-DEC-2000;	2000WO-US032678.		
PR	28-FEB-2001;	2001WO-US006520.		
PR	01-MAR-2001;	2001WO-US006866.		
PR	01-JUN-2001;	2001WO-US017800.		
PR	20-JUN-2001;	2001WO-US019692.		
PR	29-JUN-2001;	2001WO-US021066.		
PR	09-JUL-2001;	2001WO-US021735.		
PR	04-SEP-2001;	2001US-00946374.		
XX				
PA	(GETH ) GENENTECH INC.			
XX				
PI	Baker KP, Botstein D, Desnovers L, Eaton DL, Ferrara N, Fong S;			
PI	Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;			
PI	Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;			
PI	Williams PW, Wood WI;			
XX				
XX	WPI; 2003-555602/52.			
DR	P-PSDB; ADC18063.			
XX				

Mon Jun 21 09:00:25 2004

Db	841	ATCCGATGATCATGAGGACAACTGACCTGTTCTCCACCTCCACCCCTTAA	900
Qy	901	CTTGGGTACCCCTCTGGCCCTCAGAGCACCATACTCTCCATCACTTCCCTAGCTCC	960
Db	901	CTTGGGTACCCCTCTGGCCCTCAGAGCACCATACTCTCCATCACTTCCCTAGCTCC	960
Qy	961	ACTCTTCTGCGCTGGGAACTTCTTGAACTTTAACTCTCCGACGCCCTTCTAAGACCCA	1020
Db	961	ACTCTTCTGCGCTGGGAACTTCTTGAACTTTAACTCTCCGACGCCCTTCTAAGACCCA	1020
Qy	1021	CGACGGGGTGAGAGAGTGTGCAATGTCTGGAATAAATAATGAAGAGGGGAAA	1080
Db	1021	CGACGGGGTGAGAGAGTGTGCAATGTCTGGAATAAATAATGAAGAGGGGAAA	1080
Qy	1081	AAAAAAAAAAAA	1091
Db	1081	AAAAAAAAAAAA	1091

RESULT 9

ADD70708 standard; cDNA; 1091 BP.

AC ADD70708;

DT 15-JAN-2004 (first entry)

XX Human cDNA encoding secreted/transmembrane protein PRO1303.

DE Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;  
 KW immune response; cardiac insufficiency disorder; calcium flux;  
 KW umbilical vein endothelial cell; bone disorder; cartilage disorder;  
 KW arthritis; wound healing; diabetes; skeletal muscle cells; obesity;  
 KW Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;  
 KW dermatitis; herpeticiformis; Crohn's disease; thalassaemia.

OS Homo sapiens.

XX US2003099625-A1.

PN 29-MAY-2003.

PD 12-DEC-2001; 2001US-00015386.

PF 01-SEP-1998; 98US-0098716P.

PR 01-SEP-1998; 98US-0098723P.

PR 01-SEP-1998; 98US-0098749P.

PR 01-SEP-1998; 98US-0098750P.

PR 02-SEP-1998; 98US-0098803P.

PR 02-SEP-1998; 98US-0098821P.

PR 02-SEP-1998; 98US-0098843P.

PR 09-SEP-1998; 98US-0099536P.

PR 09-SEP-1998; 98US-0099596P.

PR 09-SEP-1998; 98US-0099602P.

PR 09-SEP-1998; 98US-0099642P.

PR 10-SEP-1998; 98US-0099741P.

PR 10-SEP-1998; 98US-0099754P.

PR 10-SEP-1998; 98US-0099763P.

PR 10-SEP-1998; 98US-0099792P.

PR 10-SEP-1998; 98US-0099808P.

PR 10-SEP-1998; 98US-0099812P.

PR 10-SEP-1998; 98US-0099815P.

PR 10-SEP-1998; 98US-0099816P.

PR 15-SEP-1998; 98US-0100385P.

PR 15-SEP-1998; 98US-0100388P.

PR 15-SEP-1998; 98US-0100390P.

PR 16-SEP-1998; 98US-0100584P.

PR 16-SEP-1998; 98US-0100627P.

PR 16-SEP-1998; 98US-0100661P.

PR	17-SEP-1998;	98US-0100683P.
PR	17-SEP-1998;	98US-0100684P.
PR	17-SEP-1998;	98US-0100710P.
PR	17-SEP-1998;	98US-0100711P.
PR	17-SEP-1998;	98US-0100919P.
PR	17-SEP-1998;	98US-0100930P.
PR	18-SEP-1998;	98US-0100848P.
PR	18-SEP-1998;	98US-0100849P.
PR	18-SEP-1998;	98US-0101014P.
PR	18-SEP-1998;	98US-0101068P.
PR	18-SEP-1998;	98US-0101071P.
PR	22-SEP-1998;	98US-0101279P.
PR	23-SEP-1998;	98US-0101471P.
PR	23-SEP-1998;	98US-0101472P.
PR	23-SEP-1998;	98US-0101474P.
PR	23-SEP-1998;	98US-0101475P.
PR	23-SEP-1998;	98US-0101476P.
PR	23-SEP-1998;	98US-0101477P.
PR	24-SEP-1998;	98US-0101479P.
PR	24-SEP-1998;	98US-0101738P.
PR	24-SEP-1998;	98US-0101741P.
PR	24-SEP-1998;	98US-0101743P.
PR	24-SEP-1998;	98US-0101915P.
PR	24-SEP-1998;	98US-0101916P.
PR	29-SEP-1998;	98US-0102240P.
PR	29-SEP-1998;	98US-0102240P.
PR	29-SEP-1998;	98US-0102307P.
PR	29-SEP-1998;	98US-0102330P.
PR	29-SEP-1998;	98US-0102331P.
PR	30-SEP-1998;	98US-0102484P.
PR	30-SEP-1998;	98US-0102487P.
PR	30-SEP-1998;	98US-0102570P.
PR	30-SEP-1998;	98US-0102571P.
PR	01-OCT-1998;	98US-0102684P.
PR	01-OCT-1998;	98US-0102687P.
PR	02-OCT-1998;	98US-0102965P.
PR	06-OCT-1998;	98US-0103258P.
PR	06-OCT-1998;	98US-0103449P.
PR	07-OCT-1998;	98US-0103314P.
PR	07-OCT-1998;	98US-0103315P.
PR	07-OCT-1998;	98US-0103328P.
PR	07-OCT-1998;	98US-0103393P.
PR	07-OCT-1998;	98US-0103396P.
PR	07-OCT-1998;	98US-0103401P.
PR	08-OCT-1998;	98US-0103633P.
PR	08-OCT-1998;	98US-0103678P.
PR	08-OCT-1998;	98US-0103679P.
PR	08-OCT-1998;	98US-0103711P.
PR	14-OCT-1998;	98US-0104257P.
PR	20-OCT-1998;	98US-0104987P.
PR	20-OCT-1998;	98US-0105000P.
PR	20-OCT-1998;	98US-0105002P.
PR	21-OCT-1998;	98US-0105104P.
PR	22-OCT-1998;	98US-0105169P.
PR	22-OCT-1998;	98US-0105266P.
PR	26-OCT-1998;	98US-0105693P.
PR	26-OCT-1998;	98US-0105694P.
PR	27-OCT-1998;	98US-0105807P.
PR	27-OCT-1998;	98US-0105881P.
PR	27-OCT-1998;	98US-0105882P.
PR	27-OCT-1998;	98US-0106062P.
PR	28-OCT-1998;	98US-0106023P.
PR	28-OCT-1998;	98US-0106029P.
PR	28-OCT-1998;	98US-0106030P.
PR	28-OCT-1998;	98US-0106032P.
PR	28-OCT-1998;	98US-0106033P.
PR	28-OCT-1998;	98US-0106178P.
PR	29-OCT-1998;	98US-0106248P.
PR	29-OCT-1998;	98US-0106384P.
PR	29-OCT-1998;	98US-0108500P.
PR	30-OCT-1998;	98US-0106464P.
PR	03-NOV-1998;	98US-0106856P.
PR	03-NOV-1998;	98US-0106902P.

PS	Claim 2; SEQ ID NO 193; 553pp; English.
XX	The invention relates to an isolated PRO polypeptide (secreted or
CC	Query Match 100.0%; Score 1091; DB 9; Length 1091;
	Best Local Similarity 100.0%; Pred. No. 1.5e-247;
	Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CAAGCAGGTGATCCCTTGTTGACCTTCAAGAGACAGAGAGGCGAGAGGTGGGGGC 60
Db	1 CAAGCAGGTGATCCCTTGTTGACCTTCAAGAGAGACAGAGAGGCGAGAGGTGGGGGC 60
QY	61 ACAGGAAAGGGTGACCTCTGAGATTCCCTTTTCCCCAGACATTTCGGAAGTGCACCCACC 120
Db	61 ACAGGAAAGGGTGACCTCTGAGATTCCCTTTTCCCCAGACATTTCGGAAGTGCACCCACC 120
QY	121 ATGGGGCTCAGCATCTTTTTCCTCTGTGTGTTCTTGGGCTCAGCCAGGCGACACACCG 180
Db	121 ATGGGGCTCAGCATCTTTTTCCTCTGTGTGTTCTTGGGCTCAGCCAGGCGACACACCG 180
QY	181 AAGATTTCATGSCACTGAGTGTGGCGCTAACTCAGACCGTGGCAGGTGGGGCTGTTT 240
Db	181 AAGATTTCATGSCACTGAGTGTGGCGCTAACTCAGACCGTGGCAGGTGGGGCTGTTT 240
QY	241 GAGGSCACCAAGCCTGCGCTCGGGGGTGTCTTTATGACCA CAGGTGGGTCTCTACACGC 300
Db	241 GAGGSCACCAAGCCTGCGCTCGGGGGTGTCTTTATGACCA CAGGTGGGTCTCTACACGC 300
QY	301 GCTCAGTCAGCGCAGCAGTACTGGTGGCGCTGGGGGAACACAGCCTCAGCCAGCTC 360
Db	301 GCTCAGTCAGCGCAGCAGTACTGGTGGCGCTGGGGGAACACAGCCTCAGCCAGCTC 360
QY	361 GACTGACCGAGAGATCCGSCACAGCGGCTTCTCTGTGACCCATCCCGGTACTCGGA 420
Db	361 GACTGACCGAGAGATCCGSCACAGCGGCTTCTCTGTGACCCATCCCGGTACTCGGA 420
QY	421 GCCTCGAGAGCCAGAGCAGCAGCTCGGCTGTGTGGGCTTGGCGCTCGCGCTCGCGCTA 480
Db	421 GCCTCGAGAGCCAGAGCAGCAGCTCGGCTGTGTGGGCTTGGCGCTCGCGCTCGCGCTA 480
QY	481 ACCAGCAGCTTCAACCCCTGCCCTGCGCAATGACTGTGCAACCGCTGGCAGCCAGTGC 540
Db	481 ACCAGCAGCTTCAACCCCTGCCCTGCGCAATGACTGTGCAACCGCTGGCAGCCAGTGC 540
QY	541 CACGTCACAGCTGGGGCATCAACCAACACCGGATGCTGCAACCGCTGGCAGCCAGTGC 600
Db	541 CACGTCACAGCTGGGGCATCAACCAACACCGGATGCTGCAACCGCTGGCAGCCAGTGC 600
QY	601 TGCCTCAACCTCCATCGTCTCCATGCGACCTGCCATGCTGTGTATCCCGGGAGAAATC 660
Db	601 TGCCTCAACCTCCATCGTCTCCATGCGACCTGCCATGCTGTGTATCCCGGGAGAAATC 660
QY	661 ACGAGCAACATGGTGTGTCAGCGCGGCTCCCGGGCAGGATGCTTGCAGGGTGAATCT 720
Db	661 ACGAGCAACATGGTGTGTCAGCGCGGCTCCCGGGCAGGATGCTTGCAGGGTGAATCT 720
QY	721 GGGGGCCCCCTGGTGTGCGGGGAGTCTTCAAGGTCTGTGTCTTGGGGTCTGTGGGG 780
Db	721 GGGGGCCCCCTGGTGTGCGGGGAGTCTTCAAGGTCTGTGTCTTGGGGTCTGTGGGG 780
QY	781 CCCTGTGACAGATGGCATCCCTGGAGTCTACCTATATTGCAAGTATGTGGACTGG 840
Db	781 CCCTGTGACAGATGGCATCCCTGGAGTCTACCTATATTGCAAGTATGTGGACTGG 840
QY	841 ATCCGGATGATCATGAGGAACAACTGACCTGTTTCTCCACTCCACCCGACCCCTTAA 900
Db	841 ATCCGGATGATCATGAGGAACAACTGACCTGTTTCTCCACTCCACCCGACCCCTTAA 900
QY	901 CTTGGGTACCCCTCTGGGCCCTCAGAGCAACCAATATCTCCTCCATCACTTCCCTAGCTCC 960
Db	901 CTTGGGTACCCCTCTGGGCCCTCAGAGCAACCAATATCTCCTCCATCACTTCCCTAGCTCC 960
QY	961 ACTCTGTTGGGCTGGGAACCTTCTTGGAACTTTTAACTCCTGCGAGCCCTTCTAAGACCCA 1020



Db	961	ACTCTTGTTGGCTGGGAACCTCTTGGAACTTTAACTCTCGCCAGCCCTTCTAAGACCCA	1020
Qy	1021	CGAGCGGGTGAGAGAAGTGTGCAATAGTCTCGAATAAATAAATGAAGGAGGGGCAAA	1080
Db	1021	CGAGCGGGTGAGAGAAGTGTGCAATAGTCTCGAATAAATAAATGAAGGAGGGGCAAA	1080
Qy	1081	AAAAAATAAAAA 1091	
Db	1081	AAAAAATAAAAA 1091	
RESULT 10			
ADD39785			
ID	ADD39785	standard; cDNA; 1091 BP.	
XX	XX	AC	
XX	XX	ADD39785;	
XX	XX	15-JAN-2004 (first entry)	
XX	XX	Human cDNA encoding secreted/transmembrane protein PRO1303.	
DE	DE	Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;	
KW	KW	immune response; cardiac insufficiency disorder; calcium flux;	
KW	KW	umbilical vein endothelial cell; bone disorder; cartilage disorder;	
KW	KW	arthritis; wound healing; diabetes; skeletal muscle cells; obesity;	
KW	KW	Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;	
KW	KW	dermatitis; herpeticiformis; Crohn's disease; thalassaemia.	
OS	OS	Homo sapiens.	
XX	XX	US2003083462-A1.	
XX	XX	01-MAY-2003.	
XX	XX	10-DEC-2001; 2001US-00013913.	
XX	XX	05-JAN-1999; 99WO-US000106.	
PR	PR	01-SEP-1999; 99WO-US020111.	
PR	PR	15-SEP-1999; 99WO-US021194.	
PR	PR	30-NOV-1999; 99WO-US028313.	
PR	PR	02-DEC-1999; 99WO-US028551.	
PR	PR	16-DEC-1999; 99WO-US030095.	
PR	PR	05-JAN-2000; 2000WO-US000219.	
PR	PR	06-JAN-2000; 2000WO-US000376.	
PR	PR	11-FEB-2000; 2000WO-US003565.	
PR	PR	18-FEB-2000; 2000WO-US004342.	
PR	PR	24-FEB-2000; 2000WO-US005004.	
PR	PR	02-MAR-2000; 2000WO-US005841.	
PR	PR	15-MAR-2000; 2000WO-US006884.	
PR	PR	17-MAY-2000; 2000WO-US013705.	
PR	PR	22-MAY-2000; 2000WO-US014042.	
PR	PR	30-MAY-2000; 2000WO-US014941.	
PR	PR	02-JUN-2000; 2000WO-US015264.	
PR	PR	23-AUG-2000; 2000WO-US023522.	
PR	PR	24-AUG-2000; 2000WO-US023328.	
PR	PR	08-NOV-2000; 2000WO-US030952.	
PR	PR	10-NOV-2000; 2000WO-US030873.	
PR	PR	01-DEC-2000; 2000WO-US032678.	
PR	PR	28-FEB-2001; 2001WO-US006520.	
PR	PR	01-MAR-2001; 2001WO-US006666.	
PR	PR	01-JUN-2001; 2001WO-US017800.	
PR	PR	20-JUN-2001; 2001WO-US019692.	
PR	PR	29-JUN-2001; 2001WO-US021066.	
PR	PR	04-JUL-2001; 2001WO-US021735.	
PR	PR	09-SEP-2001; 2001US-00946374.	

(GETH ) GENENTECH INC.

XX	Baker KP,	Botstein D,	Desnovers L,	Eaton D <sup>1</sup> ,	Ferrara N,	Fong S;
XX	PA	Gao W,	Goddard A,	Godowski PJ,	Grimaldi JC,	Gurney AL,
XX	PPI	Pan J,	Paoni NF,	Roy MA,	Smith V,	Stewart TA,
PPI	PPI	Williams PM,	Wood WI;			Tomas D,
PPI						Watanabe CK;

WPI; 2003-755122/71.  
P-ESDB; ADD39786.

New secreted and transmembrane PRO polypeptides useful for treating cancers, kidney disorders, Crohn's disease, diabetes mellitus, hyper- or hypo-insulinemia, sports injuries and arthritis.

Claim 2; SEQ ID NO 193; 557pp; English.

The invention relates to an isolated PRO polypeptide (secreted or transmembrane protein) having at least 80% amino acid sequence identity to an amino acid sequence chosen from 123 fully defined sequences as given in the specification (including their extracellular domains either with or without their associated signal peptides. Also include are the nucleotide (NA) sequences encoding PRO, a vector comprising the PRO NA, a host cell comprising the vector, producing PRO, a chimeric molecule comprising PRO fused to a heterologous amino acid sequence, and an anti-PRO antibody. PRO is useful as molecular weight markers for protein electrophoresis and also for chromosome identification. PRO is also useful for tissue typing. PRO and PRO NA are useful as hybridisation probes for a cDNA library to isolate the full-length PRO cDNA. PRO NA is useful for generating transgenic animals or knock-out animals which are useful in development and screening useful reagents. PRO NA is also useful in gene therapy. PRO1244, PRO1286 and PRO1303 polypeptides are useful for treating cancerous tumours. PRO1250, PRO1418 and PRO1410 polypeptides are useful for suppressing immune response. PRO1246 polypeptide is useful for treating cardiac insufficiency disorders. PRO1246 polypeptide is also useful for stimulating calcium flux in human PRO1561 polypeptide are useful for stimulating calcium flux in human umbilical vein endothelial cells. PRO1265, PRO1250 and PRO1474 polypeptides are useful for treating bone and/or cartilage disorders (e.g., arthritis) and wound healing. PRO1130, PRO1275 and PRO1418 polypeptides are useful for treating diabetes in skeletal muscle cells and obesity. PRO1265, PRO1244 and PRO1382 polypeptides are useful for treating Berger disease or other nephropathies associated with Schönlein-Henoch purpura, coeliac disease, dermatitis, herpeticiformis or Crohn's disease. PRO1478, PRO1265, PRO1412, PRO1279, PRO1304, PRO1306, PRO1418, PRO1410 and PRO1575 are useful in treating thalassaemia. The present invention also provides a PRO protein of the invention.

XX	Sequence	1091 BP;	221 A;	335 C;	313 G;	222 T;	0 U;	0 Other;
50	Query Match	100.0%;	Score 1091;	DB 9;	Length 1091;			
	Best Local Similarity	100.0%;	Pred. No. 1.5e-247;					
	Matches 1091;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps
20y	1	CAACGAGGT	CATCCCTCTGGT	GACCTTCA	AAGAGAGAGAGAGAGG	CGCAGAGGTGGGGGC	60	
Db	1	CACGAGGT	CATCCCTTGGT	GACCTTCA	AAGAGAGAGAGAGAGG	CGCAGAGGTGGGGGC	60	
20y	61	ACAGGGA	AAAGGTGACCTCTG	AGATTCCCTTTCC	CCGAGACTTTG	AGAGTACCCACC	120	
Db	61	ACAGGGA	AAAGGTGACCTCTG	AGATTCCCTTTCC	CCGAGACTTTG	AGAGTACCCACC	120	
20y	121	ATGGGGCT	CAGCATCTTTTGT	CTCTGTGTCTT	TGGCTCAG	CCAGGAGCCACACG	180	
Db	121	ATGGGGCT	CAGCATCTTTTGT	CTCTGTGTCTT	TGGCTCAG	CCAGGAGCCACACG	180	
20y	181	AAGATTTT	CAATGGCACTG	AGTGTGGCGT	TAACTACAC	CGCTGCGAGTGGGGCTGTTT	240	
Db	181	AAGATTTT	CAATGGCACTG	AGTGTGGCGT	TAACTACAC	CGCTGCGAGTGGGGCTGTTT	240	
20y	241	GAGGGCA	CCAGCCTG	CGCTCGGGGGT	GTCTTATT	TGACCAAGTGTGGTCTCTACACG	300	
Db	241	GAGGGCA	CCAGCCTG	CGCTCGGGGGT	GTCTTATT	TGACCAAGTGTGGTCTCTACACG	300	
20y	301	GCTCAT	GTGACGGC	ACAGGTACTGG	TGGCGCTCGGGG	MAACAAGCTCAGCCAGATC	360	
Db	301	GCTCAT	GTGACGGC	ACAGGTACTGG	TGGCGCTCGGGG	MAACAAGCTCAGCCAGATC	360	
20y	361	GACTGG	ACCGAGCAGATCC	CGGCACAG	CGGGTCTCTGT	GACCCATCCCGGCTACTCTGGGA	420	





[illegible]



Mon Jun 21 09:00:25 2004

Db	601	TGCTCAACCTCTCCATCGCTCCCATGCCACCTGCCATGGTGTGTATCCCGGAGAAATC	660	PR	10-SEP-1998;	98US-0099754P.
Qy	661	ACGAGCAACATGTTGTGTGAGCGGGCTCCCGGGGAGGATCCCTGCCAGGTTGATTCT	720	PR	10-SEP-1998;	98US-0099763P.
Db	661	ACGAGCAACATGTTGTGTGAGCGGGCTCCCGGGGAGGATCCCTGCCAGGTTGATTCT	720	PR	10-SEP-1998;	98US-0099792P.
Qy	721	GGGGCCCCCTGGTGTGTGGGGAGTCCCTCAAGGTCGTGTGCTGGGGGCTGTGGGG	780	PR	10-SEP-1998;	98US-0099808P.
Db	721	GGGGCCCCCTGGTGTGTGGGGAGTCCCTCAAGGTCGTGTGCTGGGGGCTGTGGGG	780	PR	10-SEP-1998;	98US-0099812P.
Qy	781	CCCTGTGGCAAGATGCATCCCTGGAGTCTACACCTATATTTCGAAGTATGTGGACTGG	840	PR	10-SEP-1998;	98US-0099815P.
Db	781	CCCTGTGGCAAGATGCATCCCTGGAGTCTACACCTATATTTCGAAGTATGTGGACTGG	840	PR	10-SEP-1998;	98US-0099816P.
Qy	841	ATCCGGATGATCATGAGAAACAATGACCTGTTTCCTCCACCTCCACCCCAACCCCTTAA	900	PR	10-SEP-1998;	98US-0100385P.
Db	841	ATCCGGATGATCATGAGAAACAATGACCTGTTTCCTCCACCTCCACCCCAACCCCTTAA	900	PR	10-SEP-1998;	98US-0100388P.
Qy	901	CTTGGATACCCCTCTGGCCCTCAGAGCAACAAATATCTCTCCATCATTCCCTAGCTCC	960	PR	10-SEP-1998;	98US-0100390P.
Db	901	CTTGGATACCCCTCTGGCCCTCAGAGCAACAAATATCTCTCCATCATTCCCTAGCTCC	960	PR	10-SEP-1998;	98US-0100642P.
Qy	961	ACTCTGTGGCTGGGAACTTCTTGGAACTTTAACTCTGCCAGCCCTTCTTAAGACCCA	1020	PR	10-SEP-1998;	98US-0100644P.
Db	961	ACTCTGTGGCTGGGAACTTCTTGGAACTTTAACTCTGCCAGCCCTTCTTAAGACCCA	1020	PR	10-SEP-1998;	98US-0100646P.
Qy	1021	CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAAATAATAATAATGAAGGGGGGCAAA	1080	PR	10-SEP-1998;	98US-0100647P.
Db	1021	CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAAATAATAATAATGAAGGGGGGCAAA	1080	PR	10-SEP-1998;	98US-0100648P.
Qy	1081	AAAAAAAAAA 1091		PR	10-SEP-1998;	98US-0100649P.
Db	1081	AAAAAAAAAA 1091		PR	10-SEP-1998;	98US-0100650P.
RESULT 13						98US-0100651P.
ID	ADD39308 standard; cDNA; 1091 BP.					98US-0100652P.
XX	AC					98US-0100653P.
XX	ADD39308;					98US-0100654P.
DT	15-JAN-2004 (first entry)					98US-0100655P.
XX	Human cDNA encoding secreted/transmembrane protein PRO1303.					98US-0100656P.
XX	Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;					98US-0100657P.
KW	immune response; cardiac insufficiency disorder; calcium flux;					98US-0100658P.
KW	umbilical vein endothelial cell; bone disorder; cartilage disorder;					98US-0100659P.
KW	arthritis; wound healing; diabetes; skeletal muscle cells; obesity;					98US-0100660P.
KW	Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;					98US-0100661P.
KW	dermatitis; herpeticiformis; Crohn's disease; thalassemia.					98US-0100662P.
XX	Homo sapiens.					98US-0100663P.
XX	US2003096954-A1.					98US-0100664P.
PN	22-MAY-2003.					98US-0100665P.
PD	07-DEC-2001; 2001US-00011671.					98US-0100666P.
PF	01-SEP-1998; 98US-0098716P.					98US-0100667P.
XX	01-SEP-1998; 98US-0098723P.					98US-0100668P.
PR	01-SEP-1998; 98US-0098749P.					98US-0100669P.
PR	01-SEP-1998; 98US-0098750P.					98US-0100670P.
PR	02-SEP-1998; 98US-0098803P.					98US-0100671P.
PR	02-SEP-1998; 98US-0098821P.					98US-0100672P.
PR	02-SEP-1998; 98US-0098843P.					98US-0100673P.
PR	03-SEP-1998; 98US-0099536P.					98US-0100674P.
PR	09-SEP-1998; 98US-0099596P.					98US-0100675P.
PR	09-SEP-1998; 98US-0099598P.					98US-0100676P.
PR	09-SEP-1998; 98US-0099602P.					98US-0100677P.
PR	09-SEP-1998; 98US-0099642P.					98US-0100678P.
PR	10-SEP-1998; 98US-0099741P.					98US-0100679P.

PR	27-OCT-1998;	98US-0105881P.	XX	(GETH ) GENENTECH INC.	
PR	27-OCT-1998;	98US-0105882P.	PA	Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;	
PR	27-OCT-1998;	98US-0106062P.	XX	Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;	
PR	28-OCT-1998;	98US-0106023P.	PI	Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;	
PR	28-OCT-1998;	98US-0106029P.	PI	Williams PM, Wood WI;	
PR	28-OCT-1998;	98US-0106030P.	PI		
PR	28-OCT-1998;	98US-0106032P.	XX	WPI; 2003-786999/74.	
PR	28-OCT-1998;	98US-0106033P.	DR	P-PSDB; ADD19309.	
PR	28-OCT-1998;	98US-0106178P.	DR		
PR	29-OCT-1998;	98US-0106248P.	XX		
PR	29-OCT-1998;	98US-0106384P.	XX		
PR	29-OCT-1998;	98US-0108500P.	PT	Novel isolated PRO polypeptide useful for tissue typing, modulating	
PR	29-OCT-1998;	98US-0106464P.	PT	biological activity of cell, as molecular weight markers in protein	
PR	30-OCT-1998;	98US-0106464P.	PT	electrophoresis, for treating arthritis, tumor.	
PR	03-NOV-1998;	98US-0106856P.	XX	Claim 2; SEQ ID NO 193; 550pp; English.	
PR	03-NOV-1998;	98US-0106902P.	PS	The invention relates to an isolated PRO polypeptide (secreted or	
PR	03-NOV-1998;	98US-0106905P.	XX		
PR	03-NOV-1998;	98US-0106919P.	XX		
PR	03-NOV-1998;	98US-0106932P.	XX		
PR	03-NOV-1998;	98US-0106934P.	XX		
PR	10-NOV-1998;	98US-0107783P.	XX		
PR	17-NOV-1998;	98US-0108775P.	XX		
PR	17-NOV-1998;	98US-0108775P.	XX		
PR	17-NOV-1998;	98US-0108787P.	XX		
PR	17-NOV-1998;	98US-0108788P.	XX		
PR	17-NOV-1998;	98US-0108801P.	XX		
PR	17-NOV-1998;	98US-0108802P.	XX		
PR	17-NOV-1998;	98US-0108806P.	XX		
PR	17-NOV-1998;	98US-0108807P.	XX		
PR	17-NOV-1998;	98US-0108867P.	XX		
PR	17-NOV-1998;	98US-0108925P.	XX		
PR	18-NOV-1998;	98US-0108848P.	XX		
PR	18-NOV-1998;	98US-0108849P.	XX		
PR	18-NOV-1998;	98US-0108850P.	XX		
PR	18-NOV-1998;	98US-0108851P.	XX		
PR	18-NOV-1998;	98US-0108852P.	XX		
PR	18-NOV-1998;	98US-0108858P.	XX		
PR	18-NOV-1998;	98US-0108904P.	XX		
PR	22-DEC-1998;	98US-0113296P.	XX		
PR	30-DEC-1998;	98US-0114233P.	XX		
PR	05-JAN-1999;	99WO-US000106.	XX		
PR	16-APR-1999;	99US-0129674P.	XX		
PR	23-JUN-1999;	99US-0141037P.	XX		
PR	20-JUN-1999;	99US-0144758P.	XX		
PR	26-JUL-1999;	99US-0145638P.	XX		
PR	01-SEP-1999;	99WO-US020111.	XX		
PR	15-SEP-1999;	99WO-US021194.	XX		
PR	29-OCT-1999;	99US-0162506P.	XX		
PR	30-NOV-1999;	99WO-US028313.	XX		
PR	02-DEC-1999;	99WO-US028551.	XX		
PR	16-DEC-1999;	99WO-US030095.	XX		
PR	05-JAN-2000;	2000WO-US000219.	XX		
PR	06-JAN-2000;	2000WO-US000376.	XX		
PR	11-FEB-2000;	2000WO-US003565.	XX		
PR	18-FEB-2000;	2000WO-US004342.	XX		
PR	24-FEB-2000;	2000WO-US005004.	XX		
PR	02-MAR-2000;	2000WO-US005841.	XX		
PR	15-MAR-2000;	2000WO-US006884.	XX		
PR	17-MAY-2000;	2000WO-US013705.	XX		
PR	22-MAY-2000;	2000WO-US014042.	XX		
PR	30-MAY-2000;	2000WO-US014941.	XX		
PR	02-JUN-2000;	2000WO-US015264.	XX		
PR	23-AUG-2000;	2000WO-US023532.	XX		
PR	24-AUG-2000;	2000WO-US023328.	XX		
PR	08-NOV-2000;	2000WO-US030952.	XX		
PR	10-NOV-2000;	2000WO-US030873.	XX		
PR	01-DEC-2000;	2000WO-US032678.	XX		
PR	28-FEB-2001;	2001WO-US006520.	XX		
PR	01-MAR-2001;	2001WO-US006666.	XX		
PR	01-JUN-2001;	2001WO-US017800.	XX		
PR	20-JUN-2001;	2001WO-US019692.	XX		
PR	29-JUN-2001;	2001WO-US021066.	XX		
PR	09-JUL-2001;	2001WO-US021735.	XX		
PR	04-SEP-2001;	2001US-00946374.	XX		

Db 721 GGGGGCCCCCTGGTGTGTGGGGAGTCTTCAAGGTCTGTGTCTGGGGGTCTGTGGGG 780  
Qy 781 CCCTGTGACAAAGATGGATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGGACTGG 840  
Db 781 CCCTGTGACAAAGATGGATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGGACTGG 840  
Qy 841 ATCCGGATGATCATGAGAAACAAGTACCTGTTTCTCCACCTCCACCCCAACCCCTTAA 900  
Db 841 ATCCGGATGATCATGAGAAACAAGTACCTGTTTCTCCACCTCCACCCCAACCCCTTAA 900  
Qy 901 CTTGGGTACCCCTCTGGCCCTCAGAGCAACCAATATCTCTCATCATCTCCCTAGCTCC 960  
Db 901 CTTGGGTACCCCTCTGGCCCTCAGAGCAACCAATATCTCTCATCATCTCCCTAGCTCC 960  
Qy 961 ACTCTGTGTGGCTGGGACCTCTTGGAACTTTTAACTCTCTCCAGCCCTTCTAAGACCCA 1020  
Db 961 ACTCTGTGTGGCTGGGACCTTTTGGAACTTTTAACTCTCTCCAGCCCTTCTAAGACCCA 1020  
Qy 1021 CGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAAATAATAATGAAGAGGGGCAAA 1080  
Db 1021 CGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAAATAATAATGAAGAGGGGCAAA 1080  
Qy 1081 AAAAAAAAAA 1091  
Db 1081 AAAAAAAAAA 1091  
RESULT 14  
ADD38831  
ID ADD38831 standard; cDNA; 1091 BP.  
XX  
AC ADD38831;  
XX  
DT 15-JAN-2004 (first entry)  
DE Human cDNA encoding secreted/transmembrane protein PRO1303.  
XX  
KW Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;  
KW immune response; cardiac insufficiency disorder; calcium flux;  
KW umbilical vein endothelial cell; bone disorder; cartilage disorder;  
KW arthritis; wound healing; diabetes; skeletal muscle cells; obesity;  
KW Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;  
KW dermatitis; herpeticiformis; Crohn's disease; thalassaemia.  
XX  
OS Homo sapiens.  
XX  
PN US2003092061-A1.  
XX  
PD 15-MAY-2003.  
XX  
PF 06-DEC-2001; 2001US-00007194.  
XX  
PX 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 01-SEP-1998; 98US-0098749P.  
PR 01-SEP-1998; 98US-0098750P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099536P.  
PR 09-SEP-1998; 98US-0099596P.  
PR 09-SEP-1998; 98US-0099598P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 09-SEP-1998; 98US-0099642P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099792P.  
PR 10-SEP-1998; 98US-0099808P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 10-SEP-1998; 98US-0099815P.  
PR 10-SEP-1998; 98US-0099816P.  
PR 15-SEP-1998; 98US-0100385P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 15-SEP-1998; 98US-0100390P.  
PR 16-SEP-1998; 98US-0100584P.  
PR 16-SEP-1998; 98US-0100627P.  
PR 16-SEP-1998; 98US-0100661P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100710P.  
PR 17-SEP-1998; 98US-0100711P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100848P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 18-SEP-1998; 98US-0101071P.  
PR 22-SEP-1998; 98US-0101279P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101474P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101476P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101741P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101915P.  
PR 24-SEP-1998; 98US-0101916P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102307P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102484P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-0103314P.  
PR 07-OCT-1998; 98US-0103315P.  
PR 07-OCT-1998; 98US-0103328P.  
PR 07-OCT-1998; 98US-0103395P.  
PR 07-OCT-1998; 98US-0103396P.  
PR 07-OCT-1998; 98US-0103401P.  
PR 08-OCT-1998; 98US-0103633P.  
PR 08-OCT-1998; 98US-0103678P.  
PR 08-OCT-1998; 98US-0103679P.  
PR 08-OCT-1998; 98US-0103711P.  
PR 14-OCT-1998; 98US-0104257P.  
PR 20-OCT-1998; 98US-0104987P.  
PR 20-OCT-1998; 98US-0105000P.  
PR 20-OCT-1998; 98US-0105002P.  
PR 21-OCT-1998; 98US-0105104P.  
PR 22-OCT-1998; 98US-0105169P.  
PR 22-OCT-1998; 98US-0105266P.  
PR 26-OCT-1998; 98US-0105693P.  
PR 26-OCT-1998; 98US-0105694P.  
PR 27-OCT-1998; 98US-0105807P.  
PR 27-OCT-1998; 98US-0105881P.  
PR 27-OCT-1998; 98US-0105882P.  
PR 27-OCT-1998; 98US-0106023P.  
PR 28-OCT-1998; 98US-0106029P.  
PR 28-OCT-1998; 98US-0106030P.  
PR 28-OCT-1998; 98US-0106032P.  
PR 28-OCT-1998; 98US-0106033P.



28-OCT-1998; 98US-0106178P.  
29-OCT-1998; 98US-0106248P.  
29-OCT-1998; 98US-0106384P.  
23-OCT-1998; 98US-0108500P.  
30-OCT-1998; 98US-0106464P.  
03-NOV-1998; 98US-0106856P.  
03-NOV-1998; 98US-0106902P.  
03-NOV-1998; 98US-0106905P.  
03-NOV-1998; 98US-0106919P.  
03-NOV-1998; 98US-0106932P.  
03-NOV-1998; 98US-0106934P.  
10-NOV-1998; 98US-0107783P.  
17-NOV-1998; 98US-0108775P.  
17-NOV-1998; 98US-0108779P.  
17-NOV-1998; 98US-0108787P.  
17-NOV-1998; 98US-0108788P.  
17-NOV-1998; 98US-0108801P.  
17-NOV-1998; 98US-0108802P.  
17-NOV-1998; 98US-0108806P.  
17-NOV-1998; 98US-0108807P.  
17-NOV-1998; 98US-0108857P.  
17-NOV-1998; 98US-0108925P.  
18-NOV-1998; 98US-0108848P.  
18-NOV-1998; 98US-0108849P.  
18-NOV-1998; 98US-0108850P.  
18-NOV-1998; 98US-0108851P.  
18-NOV-1998; 98US-0108852P.  
18-NOV-1998; 98US-0108858P.  
18-NOV-1998; 98US-0108904P.  
22-DEC-1998; 98US-0113296P.  
30-DEC-1998; 98US-0114223P.  
05-JAN-1999; 99WO-US000106.  
16-APR-1999; 99US-0129674P.  
23-JUN-1999; 99US-0141037P.  
20-JUL-1999; 99US-0144758P.  
26-JUL-1999; 99US-0145698P.  
01-SEP-1999; 99WO-US020111.  
15-SEP-1999; 99WO-US021194.  
29-OCT-1999; 99US-0162506P.  
30-NOV-1999; 99WO-US028313.  
02-DEC-1999; 99WO-US028851.  
16-DEC-1999; 99WO-US030095.  
05-JAN-2000; 2000WO-US000219.  
06-JAN-2000; 2000WO-US000376.  
11-FEB-2000; 2000WO-US003565.  
18-FEB-2000; 2000WO-US004342.  
24-FEB-2000; 2000WO-US005004.  
02-MAR-2000; 2000WO-US005841.  
15-MAR-2000; 2000WO-US006884.  
17-MAY-2000; 2000WO-US013705.  
22-MAY-2000; 2000WO-US014042.  
30-MAY-2000; 2000WO-US014941.  
02-JUN-2000; 2000WO-US015264.  
23-AUG-2000; 2000WO-US023522.  
24-AUG-2000; 2000WO-US023328.  
08-NOV-2000; 2000WO-US030952.  
10-NOV-2000; 2000WO-US030873.  
01-DEC-2000; 2000WO-US032678.  
28-FEB-2001; 2001WO-US006520.  
01-MAR-2001; 2001WO-US006666.  
01-JUN-2001; 2001WO-US017800.  
20-JUN-2001; 2001WO-US019692.  
29-JUN-2001; 2001WO-US021066.  
09-JUL-2001; 2001WO-US021735.  
04-SEP-2001; 2001WO-US0046374.  
(GETH ) GENENTECH INC.

Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;  
Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;  
Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;  
Williams PM, Wood WI;

DR WPI; 2003-765477/72.  
DR P-PSDB; ADD38832.

XX New isolated PRO polypeptide such as PRO1560, PRO444, PRO1018, PRO1773,  
PT PRO1244, PRO1246, useful for treating cancerous tumors, cardiac  
PT insufficiency disorders, wound healing, Crohn's disease, celiac disease.  
XX Claim 2; SEQ ID NO 193; 555pp; English.

XX The invention relates to an isolated PRO polypeptide (secreted or

Query Match 100.0%; Score 1091; DB 9; Length 1091;  
Best Local Similarity 100.0%; Pred. No. 1.5e-247;  
Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCAGGTATCCCTTGGTCACTTTGGTCACTTTAAAGAGAGCAGAGGCGCAGAGTGGGGGCG 60  
DB 1 CAAGCAGGTATCCCTTGGTCACTTTGGTCACTTTAAAGAGAGCAGAGGCGCAGAGTGGGGGCG 60  
QY 61 ACAGGGAAGGGTGACCTCTGAGATTCCTTTCCCGCAGACTTTGGAGTGACCCACC 120  
DB 61 ACAGGGAAGGGTGACCTCTGAGATTCCTTTCCCGCAGACTTTGGAGTGACCCACC 120  
QY 121 ATGGGGCTCAGCATCTTTTGGTCTCTGTGTGTTCTTGGGCTCAGCAGGCGACACCG 180  
DB 121 ATGGGGCTCAGCATCTTTTGGTCTCTGTGTGTTCTTGGGCTCAGCAGGCGACACCG 180  
QY 181 AAGATTTTCAATGGCACTGAGTGTGGCGTAACTCAAGCCGTGGGCGTGGGGCTGTTT 240  
DB 181 AAGATTTTCAATGGCACTGAGTGTGGCGTAACTCAAGCCGTGGGCGTGGGGCTGTTT 240  
QY 241 GAGGGCACCAGCCTCGCGTGGGGGTGCTCTTATTGACACAGGTGGGTCTCACAGCG 300  
DB 241 GAGGGCACCAGCCTCGCGTGGGGGTGCTCTTATTGACACAGGTGGGTCTCACAGCG 300  
QY 301 GCTCACTGACGCGCAGAGTACTGGGTGCGCTGCGGGAACAACAGCTCAGCAGCTC 360  
DB 301 GCTCACTGACGCGCAGAGTACTGGGTGCGCTGCGGGAACAACAGCTCAGCAGCTC 360  
QY 361 GACTGGACCGAGAGATCCGGCAGAGCGGTCTCTGTGACCCATCCCGGTACTCGGGA 420  
DB 361 GACTGGACCGAGAGATCCGGCAGAGCGGTCTCTGTGACCCATCCCGGTACTCGGGA 420  
QY 421 GCCTCGACGAGCAGCAGCAGCTCCGGTGTGCGGCTGCGGCTGCGGCTGCGGCTG 480  
DB 421 GCCTCGACGAGCAGCAGCAGCTCCGGTGTGCGGCTGCGGCTGCGGCTGCGGCTG 480  
QY 481 ACCAGCAGCGTTCAACCCCTGCGCCTGCCAATGACCTGTGCAACCGTGGCAGAGTGC 540  
DB 481 ACCAGCAGCGTTCAACCCCTGCGCCTGCCAATGACCTGTGCAACCGTGGCAGAGTGC 540  
QY 541 CACGCTCAGGCTGGGCGATCAACACACCGAGAACCCATTCGGGATCTGCTCCAG 600  
DB 541 CACGCTCAGGCTGGGCGATCAACACACCGAGAACCCATTCGGGATCTGCTCCAG 600  
QY 601 TGCCTCAACCTCTCCATGCTCCCATGCCACCTGCCATGCTGTGTATCCCGGAGATC 660  
DB 601 TGCCTCAACCTCTCCATGCTCCCATGCCACCTGCCATGCTGTGTATCCCGGAGATC 660  
QY 661 ACAGCAACATGTTGTGTGAGGGGGGTCCCGGGGAGGATGCTGCCAGGTGATTTCT 720  
DB 661 ACAGCAACATGTTGTGTGAGGGGGGTCCCGGGGAGGATGCTGCCAGGTGATTTCT 720  
QY 721 GGGGGCCCCCTGGTGTGTGGGGGAGTCTTCAAGTCTGTGTCTCTGGGGGTCTGTGGG 780  
DB 721 GGGGGCCCCCTGGTGTGTGGGGGAGTCTTCAAGTCTGTGTCTCTGGGGGTCTGTGGG 780  
QY 781 CCCTGTGACAGATGGCATCCCTGGAGTCTACACCTATATTGCAAGTATGTGACCTGG 840  
DB 781 CCCTGTGACAGATGGCATCCCTGGAGTCTACACCTATATTGCAAGTATGTGACCTGG 840  
QY 841 ATCCGGATGATCAGAGAAACACTGACCTGTTTCTCCACCTCCACCCCAACCCCTTAA 900





Mon Jun 21 09:00:25 2004

Db	961	ACTCTTGTGGCTGGGAACCTTCTTGGAACTTTAACTCCTGCAGCCCTTCTAAGACCCA	1020
Qy	1021	CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATGAAGAGGGGCAAA	1080
Db	1021	CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATGAAGAGGGGCAAA	1080
Qy	1081	AAAAAAAAAA	1091
Db	1081	AAAAAAAAAA	1091

Search completed: June 20, 2004, 02:29:02  
Job time : 514 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 01:58:51 ; Search time 3353 Seconds  
(without alignments)  
9716.564 Million cell updates/sec

Title: US-10-015-385A-193

Perfect score: 1091

Sequence: 1 caacaggtcatcccttgg.....aggggcaaaaaaaaaa 1091

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	459	42.1	890	11 AK009217	AK009217 Mus muscu
2	443	40.6	898	13 BY709025	BY709025 BY709025
3	375	34.4	472	13 BX110362	BX110362 BX110362
4	311.8	28.6	469	9 AI394679	AI394679 tg24b03.x

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

c	5	305.2	28.0	676	13	BQ195790
	6	282.6	25.9	520	13	BX514743
c	7	248.4	24.7	545	9	AI324852
	8	268.2	24.6	506	9	AA028356
	9	257.2	23.6	449	9	AI322408
	10	250.4	23.0	1186	11	BC035385
	11	221.6	20.3	391	14	W20715
	12	216	19.8	264	12	BG230967
	13	208.4	19.1	1282	11	BC015551
	14	200.4	18.4	1295	11	AK009360
	15	200.4	18.4	1295	11	AK009720
	16	196	18.0	783	29	AY410899
	17	195.4	17.9	783	29	AY410900
	18	195.4	17.9	880	14	CB202840
	19	195.4	17.9	898	14	CB204935
	20	191.6	17.6	783	29	AY410898
	21	189.2	17.3	1269	11	AK009659
	22	188.4	17.3	852	14	CB587169
	23	182.8	16.8	678	12	BI763040
	24	180.6	16.6	826	14	CB574882
	25	180.4	16.5	1072	12	BM559782
	26	174.8	16.0	1074	12	BM559617
	27	174.2	16.0	841	12	BI818697
	28	174.2	16.0	853	12	BG720793
	29	169.2	15.5	754	12	BI653899
c	30	167.8	15.4	1051	12	BM547198
	31	166	15.2	730	14	CD497946
	32	166	15.2	762	14	CD497945
c	33	166	15.2	783	14	CD498690
	34	166	15.2	785	14	CD506723
c	35	166	15.2	787	14	CD503321
	36	166	15.2	788	14	CD506194
c	37	166	15.2	799	14	CD509135
	38	166	15.2	800	14	CD506130
c	39	166	15.2	800	14	CD509521
	40	166	15.2	809	14	CD506193
	41	166	15.2	821	14	CD506722
	42	166	15.2	824	14	CD509134
	43	166	15.2	825	14	CD503320
	44	166	15.2	828	14	CD501456
	45	166	15.2	837	14	CD505738

## ALIGNMENTS

RESULT 1  
AK009217 890 bp mRNA linear HTC 20-SEP-2003  
Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:231008B01 product:similar to KALLIKREIN 12 PRECURSOR (EC 3.4.21.-) (KALLIKREIN-LIKE PROTEIN 5) (KLK-L5) [Homo sapiens], full insert sequence.

ACCESSION AK009217 GI:12843870  
VERSION HTC; CAP trapper.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636

REFERENCE  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374



Db	719	GCCTGCA---TG	CAGCTCTGCTTAGGAGCGGCTCTACCCCTCGGAGCACCAGGTCTCC	774
QY	940	TCCATCACTTCCCT	AGCTTCACTCTGTTGGCTCGGAACTTCTTGAACTTTAACTCC	999
Db	775	TCCTCACCACCT	CAGGCTCCACCCCTGTGAGCTTAGAGCTTTGAAACCTGACCTAC	834
QY	1000	TGCCAGCGCCCTT	TAAGACCCA	1020
Db	835	ACCATGGCAGCGGAGGCA	855	
RESULT 2				
LOCUS	BY709025			
DEFINITION	BY709025 RIKEN full-length enriched, adult male tongue <i>Mus musculus</i> cDNA clone 2310008B01 5', mRNA sequence.	898 bp	mRNA	linear EST 16-DEC-2002
ACCESSION	BY709025			
VERSION	BY709025.1	GI:27120221		
KEYWORDS	EST.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 898)			
AUTHORS	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikiado, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, I., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schrim, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, B. W., Blake, J. A., Bradt, D., Brusci, V., Clothier, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, R. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Walestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wyntshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kitahara, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.			
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs			
JOURNAL	Nature	420	563-573	(2002)
MEDLINE	22354683			
PMID	12466851			
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Sushiro-cho, Tsukuba, Ibaraki 305-8565, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp URL: http://genome.gsc.riken.go.jp/ Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Hirozane, T., Horii, F., Fukuda, S., Hashizume, W., Hayashida, K., Kawai, J., Kojima, Y., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Murata, M., Nakamura, M., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.			





found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)  
 Insert Length: 1051 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 397.

# FEATURES

```

1. 469
Location/Qualifiers
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:2109677"
  /tissue_type="B-cell, chronic lymphocytic leukemia"
  /lab_host="DH10B"
  /clone_lib="NCI CGAP CLL1"
  /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTCGGAGCGCGCATGCTGTCTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
```

## ORIGIN

```

Query Match      28.6%; Score 311.8; DB 9; Length 469;
Best Local Similarity 97.3%; Pred. No. 96-50;
Matches 328; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

QY 493 CAACCCCTGCCCCCTGCCAATGACTGTGCAACCGTGGCAGCGAGTGCACGCTTCAGGC 552
Db 469 CAACCCCTGCCCCCTGCCAATGACTGTGCAACCGTGGCAGCGAGTGCACGCTTCAGGC 410
QY 553 TGGGGCATCACCACACACCGGACCCATTCGCGATCTCTCCAGTGCCTCAACCTC 612
Db 409 TGGGGCAT-CCCAACACCCCGGACCCATTCGCGATCTCTCCAGTGCCTCAACCTC 352
QY 613 TCCATGCTCTCCATGCCACCTGCGATGTGTATCCCGGAGATACAGGACACATG 672
Db 351 TCCATGCTCTCCATGCCACCTGCGATGTGTATCCCGGAGATACAGGACACATG 292
QY 673 GTGTGTGAGGGGGGTCTCCCGGGCAGGATGCTGCGAGGTGATTCGGGGGCCCTCG 732
Db 291 GTGTGTGAGGGGGGTCTCCCGGGCAGGATGCTGCGAGGTGATTCGGGGGCCCTCG 232
QY 733 GTGTGTGAGGGGGGTCTCCCGGGCAGGATGCTGCGAGGTGATTCGGGGGCCCTCG 792
Db 231 GTGTGTGAGGGGGGTCTCCCGGGCAGGATGCTGCGAGGTGATTCGGGGGCCCTCG 172
QY 793 GATGGCATCCCTGGAGTCTACACCTATATTGCAAGT 829
Db 171 GATGGCATCCCTGGAGTCTACACCTATATTGCAACT 135
```

## RESULT 5

```

BQ195790/c
LOCUS
DEFINITION
  UI-R-CN1-cmq-p-20-0-UI.s1 UI-R-CN1 Rattus norvegicus cDNA clone
  UI-R-CN1-cmq-p-20-0-UI 3', mRNA sequence.
ACCESSION
  BQ195790
VERSION
  BQ195790.1 GI:20371341
KEYWORDS
  EST.
SOURCE
  Rattus norvegicus (Norway rat)
  Rattus norvegicus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
  Rattus.
REFERENCE
  1 (bases 1 to 676)
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  JOURNAL
  Genome Res. 6 (9), 791-806 (1996)
  MEDLINE
  97044477
```

## COMMENT

8889548  
 Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. The sequence tag present in the cDNA between the NotI site  
 and the oligo-dT track served to identify it as a clone from the  
 normalized fundus library cDNA Library Preparation: M.B. Soares Lab  
 Clone distribution: clones will be available through Research  
 Genetics ([www.resgen.com](http://www.resgen.com))  
 Seq primer: M13 Forward  
 POLYA=Yes.

## FEATURES

```

source
1. 676
Location/Qualifiers
  /organism="Rattus norvegicus"
  /mol_type="mRNA"
  /strain="Sprague-Dawley"
  /db_xref="taxon:10116"
  /clone="UI-R-CN1-cmq-p-20-0-UI"
  /dev_stage="adult"
  /lab_host="DH10B (Life Technologies)"
  /clone_lib="UI-R-CN1"
  /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CN1
library is a subtracted library derived from the following
pool of seven normalized rat libraries: normalized rat
seminal vesicles, normalized rat penis, normalized rat
bladder, normalized rat cervix, normalized rat brown
adipose, normalized rat fundus, and normalized rat
salivary gland. It was constructed according to the
procedure described by Bonaldo, Lennon & Soares (Genome
Research Genome 6: 791-806, 1996). For construction of
the CN1 library, plasmid DNA from the pool of normalized
libraries was electroporated into competent bacteria for
the production of single-stranded circular DNA. This was
then used as a tracer in a subtractive hybridization with
a driver (PCR amplified inserts from a plamid DNA template
from the Rat Unigene Set corresponding to plates R-5-AA-NN
excluding plates R-5-MM and MN. This pool represented 40%
of the final driver population. b) a pool of about 29,000
clones from subtracted libraries CA0 and CA1 corresponding
to plates R-CA0-AWV through R-CA0-AXS, R-CA0-AZX through
R-CA0-BAZ, R-CA0-BFE through R-CA0-BHY, R-CA0-BJS,
R-CA0-BKE, R-CA0-BKG-H, R-CA0-BKJ-K, R-CA0-BKP through
R-CA0-BKS, R-CA0-BKU-V, R-CA0-BLY through R-CA0-BMA,
R-CA0-BMC through R-CA0-BME, R-CA0-BNS, R-CA0-BOB through
R-CA0-BOU, R-CA0-BPA through R-CA0-BPG, R-CAI-BUR,
R-CAI-BDA, R-CAI-BHZ through R-CAI-BJF, R-CAI-BKJ,
R-CAI-BKT through R-CAI-BKB, R-CAI-BKD, R-CAI-BKF,
R-CAI-BKI, R-CAI-BKT, R-CAI-BLF, R-CAI-BLH through
R-CAI-BLN, R-CAI-BLS, R-CAI-BLU-V, R-CAI-BNR, and
R-CAI-BLE. The resulting pool represented 20% of the
final driver population. c) a pool of about 15,000 clones
from non-normalized libraries CS0s, CT0s, CU0s, and CX0s
and normalized libraries CS0, CT0, CU0, CWO, and CX0
corresponding to plates R-CS0s-CBD through R-CS0s-CBO,
R-CT0s-CAM through R-CT0s-CAX, R-CU0s-CBP through
R-CU0s-CCA, R-CW0s-CCE through R-CW0s-CCM, R-CX0s-CCN
through R-CX0s-CCX, R-CS0-BSD, R-CS0-BTD through
R-CS0-BTV, R-CS0-BVM, R-CT0-BTW through R-CT0-BUP,
R-CWO-BVN, R-CWO-BUQ through R-CWO-BVL, R-CWO-BVY through
R-CWO-BWP, R-CWO-BXN through R-CWO-BXO, R-CXO-BWQ through
R-CXO-BXM. The resulting pool represented 5% of the final
driver population. d) a pool of about 5,000 clones (1,000
from non-normalized eye library CV0 and 4,000 from
normalized eye library CV1) corresponding to plates
```



TGTTACCATCTGAAGTCGAGCGCGCATTTTTTTTTTTTTTTT 3',  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified pTV73 vector  
(Pharmacia). Library went through one round of  
normalization to a Cot = 5. Library constructed by Bento  
Soares and M. Patima Bonaldi. RNA was kindly provided by  
Dr. Minoru Ku (Wayne State University). "

## ORIGIN

Query Match	25.9%;	Score 282.6;	DB 13;	Length 520;
Best Local Similarity	72.7%;	Pred. No. 3.6e-44;		
Matches 380;	Conservative	0;	Mismatches 139;	Indels 4; Gaps 1;
448	CGGCTGTCGGGCTGCGCCTGCGCCGTCGCGGTAAACAGCAGCGGTCAACCCCTGCCGCTG	507		
Dd				
2	CGACTCTCTGGGCTGAACAGACCCATCCACTGACCCGTGCTGTCGGGCCGTGGCCCTG	61		
Dd				
508	CCCAATGACTGTGCACACCGCTGGCACCGAGTGCACAGTCTCAGGCTGGGCATCAACAAC	567		
Dd				
62	CCCGATCTCTGTGAACATACAGGGGCCATGTGTCATGTCTCAGATGGGTACTTACAAC	121		
Dd				
568	CACCCAGGAACCCGATTCGCGGATCTGCTCCAGTGCTCAACCTCTCCATCGTCTCCCAT	627		
Dd				
122	AAGCCATGGACCCATTCACAGACCGGCTCCAGTGCTCAACCTCTCCATGCTCTCCAAT	181		
Dd				
628	GCACCTGCCATCGTGTGATCCCGGAGAAATCACAGCAACATGTTGTGCAGGGGGC	687		
Dd				
182	GAGACGTGCCGGGCTGTGTTCTCTGGAAGAGTACGGAGAAATATGTTGTGCAGGTGGA	241		
Dd				
688	GTCCCGGGCAGGATGCTGCCAGGGTGATTTGTGGGGGCCCCCTGGTGTGTGGGGGAGTC	747		
Dd				
242	GAAGCCGGGAAGNATGCTGTCAAGGTGACTCTGGAGGCCCTGGTGTGTGGAGGGGTT	301		
Dd				
748	CTTCAAGGTCGTGTCTCTGGGGTCTGTGGGGCCCTGTGGAAGAATGCAATCCCTGGA	807		
Dd				
302	CTTCAAGGTCGTGTTCTCGGGATCTGTTGGGCTTGTGGTCAAAAGGTATTTCCAGGA	361		
Dd				
808	GTCACACCTATATTTCAGAGTATGTGGATCGGATCGGATGATCATGAGGAACAACCTGA	867		
Dd				
362	GTCACACGAAGGTCGCAATAACACGGACTGGATCAAGATAGTCATCAGGAATATCTAA	421		
Dd				
868	CTGTGTTCTTCCACTCCACCCACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGAGC	927		
Dd				
422	GTGACTCCCTCAGCCTGGA----TGCAGCTCTGTTAGGGACCGCTCTGACCTCGAGC	477		
Dd				
928	ACCAATATCTCTCCATCACTTCCCTTAGCTTCCACTCTTGTTG	970		
Dd				
478	ACCAAGGTCCTCTCTCTCACCACCTCCAGGCTCCACCCCTCTG	520		
Dd				

RESULT	7
LOCUS	A1324852/c
DEFINITION	m20d11.x1 Soares mouse p3NMF19.5 Mus musculus cDNA clone IMAGE:464085 3' similar to SW:TRY1_CHICK_Q90627 TRYPSINOGEN P1 PRECURSOR., mRNA sequence.
ACCESSION	A1324852
VERSION	A1324852.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 545) Marra,M., Hallier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterson,R.  The WashU-HMMI Mouse EST Project Unpublished (1996) Contact: Marra M/Mouse EST Project
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

WashU-HMMI Mouse EST Project  
Washington University School of MedicineP  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
MGI:277901

This clone was previously sequenced on the 5' end only, this new data is from the 3' end  
High quality sequence stop: 475.

FEATURES	SOURCE
----------	--------

```

1. .545
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:464085"
/dev_stage="19.5 dpc total fetus"
/lab_host="VH10B (ampicillin resistant)"
/clone_lib="Soares mouse p3NM19.5"
/notice=vector: p7T73 (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTCACCAATCTGAATCGGAGCGGCGCATTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p7T73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fátima Bonaldo. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."

```

## ORIGIN

Query Match	24.7%;	Score 269.4;	DB 9;	Length 545;
Best Local Similarity	72.3%;	Pred. No. 1.2e-41;		
Matches 365;	Conservative 0;	Mismatches 136;	Indels 4;	Gaps 1;
QY	516	CTGTGCAACCGCTGGCACCGAGTGCACACGCTCTCAGGCTGGGGCATCACCAACACCCACG	575	
DB	541	CTGTGTAACATACAGGGGCCATGTGTATGTCTCAGATGGGGTACTCAACACAGCCATG	482	
QY	576	GAACCCATTCCCGGATCTGCTCCAGTGCCTTCAACCTCTCCATGCTTCCCATGCCACTG	635	
DB	481	GGACCCATTCCAGACCGGCTCCAGTGCCTCAACCTCTCCACTGTCTCCAATGAGACGTG	422	
QY	636	CGATGGTGTATTCOCGGGAGATTCACGAGCAACATGGTGTGTCAGGCGCGCTCCGGG	695	
DB	421	CCGGGCTGTCTTCTCGAAGAGTGCACGAGAAATATGTTGTGTGTCAGTGGAGAAGCGG	362	
QY	696	GCAGGATGCCGCCAGGTGATTTCTGGGGGCCCCCTGCTGTGTGGGGAGTCCTTCAAGG	755	
DB	361	GAAGGATGCCCTGTCCAGGTGACTCTGAGGCGCCCTGGTGTGTGGAGGGTCTTCTCAAGG	302	
QY	756	TCGTGCTCTCTGGGGGCTGTGTGGGGCCCTGTGGCAAGATGGCATCCCTGGAGTCTACAC	815	
DB	301	TCGTGTTTCTCTGGGATCTGTGTGGGCTTGTGGTCAAAAAGGTATTCAGGAGTCTACAC	242	
QY	816	CTATATTTGCAAGTATGTGACTGGATCCGGATGATCATGAGGAACAACTGACCTGTTTC	875	
DB	241	GAAGGTCTGCAAAATACACGGACTGGATCAGAAATAGTCAATAGGAAATCACTAAAGTGA	182	
QY	876	CTCCACCTCCACCCCCACCCCTTAACCTTGATGGTACCCCTCTGGCCCTCAGAGCACCAATAT	935	
DB	181	CTCAGCCTGCATGACG----CTCTGCTTAGGGACCGCTCTGACCCCTCGGAGCACCAAGGT	126	
QY	936	CTCCTCCATCACTTCCGCTAGCTCCACTCTTGTGTGGCTGGGAACCTTCTTGGAACTTTAA	995	
DB	125	CTCCTCCTCCACCACTCCAGGCTCCACCCTTGTGAGCTTAGAACCTCTTGAACCTCGA	66	
QY	996	CTCCTGCGAGCCCTTCTAAGACCA 1020		
DB	65	CTACACCATGGCAGCGGAGAGCA 41		

## RESULT 8

AA028356 506 bp mRNA linear EST 11-SEP-1996  
 LOCUS mi20d11.r1 Soares mouse p3NMf19.5 Mus musculus cDNA clone  
 DEFINITION IMAGE:464085 5' similar to SW:TRYP\_BOVIN P00760 TRYPSINOGEN ; mRNA sequence.

## ACCESSION

AA028356 GI:1494434

## VERSION

AA028356.1

## KEYWORDS

EST.

## SOURCE

Mus musculus (house mouse)

## ORGANISM

EST.

## REFERENCE

1 (bases 1 to 506)  
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geiseli,S., Kucaba,I., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

## TITLE

The WashU-HMI Mouse EST Project

## JOURNAL

Unpublished (1996)

## COMMENT

Contact: Marra M/Mouse EST Project  
 WashU-HMI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:277901

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 486.

## FEATURES

source

Location/Qualifiers

1..506  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:464085"  
 /dev\_stage="19.5 dpc total fetus"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares mouse p3NMf19.5"  
 /note="vector: p773D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGGCATTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p773 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Patima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."

## ORIGIN

Query Match 24.6%; Score 268.2; DB 9; Length 506;  
 Best Local Similarity 73.0%; Pred. No. 2.1e-41;  
 Matches 373; Conservative 0; Mismatches 133; Indels 5; Gaps 2;  
 QY 448 CGGCTGCTGGCGCTGGCTGCCCTGCCGCGTAAACAGCAGCGTTTCAACCCCTGCCCTG 507  
 DB 1 CCACTCCCTGGCGCTGAAACAGACCCATCCACTGACCGCTGCTGCCGCGCTG-GCCTG 59  
 QY 508 CCCAATGACTGTGCAACCGCTGGCACCAGTCCCGCTCTCAGGCTGGGGCATCCCAAC 567  
 DB 60 CCCAGTTCCTGTGTAACACAGGGCCATGTGTGTGTCTCAGGATGGGTACTACAAAC 119  
 QY 568 CACCACAGGAACCCATTCCCGGATCTGCTCCAGTCCGCTCAACCTCTCCATCGTCTCCAT 627  
 DB 120 AAGCCATGGGACCCATTCCACAGCGGCTCCAGTCCCTCAACCTCTCCACTGCTCCAAAT 179

QY 628 CCCACCTGCCATGCTGTATCCCGGAGATCAGAGCAACATGGTGTGTGCGAGCGGC 687  
 DB 180 GAGACGTGCGCGGCTGTGTTTCTTGGAGAGTACGAGAGATATGTTGTGTGCGAGTGA 239  
 QY 688 GTCCCGGGGAGGATGCTGCCAGGGTGATTTCTGGGGCCCCCTGGTGTGTGGGGAGTC 747  
 DB 240 GAAGCCGGGAGGATGCTGTACAGGTGACTCTGAGGCCCCCTGGTGTGTGGAGGGTT 299  
 QY 748 CTTCAAGGTCTGGTCTCTGGGGTCTGTGGGGCCCTGTGGACAGATGGCATCCCTGA 807  
 DB 300 CTTCAAGGTCTGGTCTCTGGGGTCTGTGGGGCCCTGTGGTCAAAAAGGATTTCCAGA 359  
 QY 808 GTCTACACTATATTGTCAGAGTATGTGACTGGATCCCGATCATCAGGAACAACATCA 867  
 DB 360 GTCTACAGAGGTCTGCAATACAGGACTGGATCAGATAGTATCATCAGGATTAATA 419  
 QY 868 CTTGTTCTCTCACTCCACCCCAACCCCTTAACCTGGTGGTACCCCTGTGGCCCTCAGAGC 927  
 DB 420 GTGACTCCCTCAGCTGCA---TGCAGTCTGTCTTAGGACCGCTCTGACCTCGGAGC 475  
 QY 928 ACCAATATCTCTCCATCACTTCCCTAGCT 958  
 DB 476 ACCAAGGTCTCTCTCTCACCACCTCAGGCT 506

## RESULT 9

AI322408

LOCUS

DEFINITION

AI322408

ACCESSION

AI322408

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AI322408

LOCUS

DEFINITION

AI322408

ACCESSION

AI322408

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AI322408

LOCUS

DEFINITION

AI322408

ACCESSION

AI322408

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AI322408

LOCUS

DEFINITION

AI322408

ACCESSION

AI322408

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT











RESULT 15  
AK009720  
LOCUS  
DEFINITION  
AK009720 1295 bp mRNA linear HTC 20-SEP-2003  
Mus musculus adult male tongue cDNA, RIKEN full-length enriched  
library, clone:2310040F07 product:protease, serine, 20, full insert  
sequence.  
AK009720  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AK009720.1 GI:12844688  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159  
3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, J.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multipipillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4  
The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 1295)  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,  
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,  
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,  
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,  
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,  
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,  
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,  
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,  
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,  
Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,  
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,

Fax: 81-45-503-9216)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
further details.  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was  
prepared by using triazole thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. Second  
strand cDNA was prepared with the primer adapter of sequence [5'  
GAGAGAGATCTCAGTTAATTAATTAATCCCCCCCCC 3']. cDNA was cleaved  
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.  
Host: SOLR.  
Location/Qualifiers  
1. 1295  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM:DB-2310040F07"  
/db\_xref="MGI:1898372"  
/db\_xref="taxon:10090"  
/clone="2310040F07"  
/sex="male"  
/tissue type="tongue"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
144..974  
/note="unnamed protein product; protease, serine, 20  
(MGI:MGI:1929977)  
putative"  
/codon\_start=1  
/protein\_id="BAB26461.2"  
/db\_xref="GI:26365928"  
/translation="MRRLKSDWLSTETREPGARPALQARMILRLIALVTHGVGG  
ETRIKGYECPHSPQWALFQKRLCCGATLAPKWLTAARCKPHYVILLGEHN  
LEKTDGCEQRMAFESPHDPNNLKNDRNDMLVKMSPVFFRAVQPLTLSPH  
CVAAGTSCILSGMGTSSPOLRLPHSLRCANVSIIEHKECEKAYGNTIDMLCASVR  
KEGKDSQGSGLVNCGLSGLQSLGWDPCAVTRKEGVYTKYCKFNWIEHVRNN  
"  
Query Match 18.4%; Score 200.4; DB 11; Length 1295;  
Best Local Similarity 57.7%; Pred. No. 2.1e-28;  
Matches 420; Conservative 0; Mismatches 296; Indels 12; Gaps 3;  
QY 179 CGAAGATTTTCAATGGCAGCTAGTGTGGCGTAACCTACAGCGTGGGAGTGGGCTGT 238  
DB 280 CGAGGATCATCAAGGGTTATGAGTGCAGGCTCACTCAGCCATGGCAGTGGCCTCT 339  
QY 239 TTGAGGGCACCAGCCTGGCGTGGGGGTGCTTATTGACACAGAGTGGTCTCTCAG 298  
DB 340 TTCAAGACAGCGCTTCTCTGTGGGCAACCTCATCGCCCAATGGCTCTTGACAG 399  
QY 299 CGGCTCTACTGCGCGGAGCAGAGTACTGGGTGCGCTGGGGGAAACACAGCCTCAGCCAGC 358  
DB 400 CAGCCCACTGCGGCAAGCCCCATTAGGTGATCTCTCTTGGAGAGACACAATCTAGAGAAGA 459  
QY 359 TCGACTGACCGAGCAGAGTCCGGCAGACGGGCTCTCTGTGACCCCATCCCGGCT----- 412  
DB 460 CAGACGGCTGTGAGCAGAGGGGATGGCCACTGAGTCTCTTCCCCCAGCCGACTTCAACA 519  
QY 413 ACCTGGGAGCTCCAGCAGCCACGAGCAGCCTCCGCTGCTGGGCTGGCGCTGGCGCG 472  
DB 520 ACAGCTCCCCCAACAAGACCACCGGAATGACATAATGCTTGTGAAGATGCTGCTCCCG 579  
QY 473 TCCGCGTAACAGCAGCGGTTTCAACCCCTGCGCCCTGCGCCCAATGACTGTGCACACCGCTGSCA 532  
DB 580 TCTTCTTTTACCGAGCTGTGAGCCACTCACTCCCTGTCCCACTGCTCGCTCGAGGCA 639  
QY 533 CCGAGTGCACAGTCTCAGGCTGGGGGATTCACCAACACCCACCGAACCCATTCGCGGATC 592

Db	640	CCAGTGCCTCATTTCTGGATGGGGCAACACGTCACGCCCCCAGTTGGCGCTGCCTCAT	699
QY	593	TGCTCCAGTGCCTCAACCTCTCCATGGTCTCCCATGCCACCTGCCATGGTGTGTATCCCG	652
Db	700	CCTTGCATGTGCCAATGTCTCCATCATCGAACACAGGAGTGTGAGAAGGCCCTACCCCG	759
QY	653	GGAGATCAGAGCAACATGGTGTGTGCGAGCG---GCGTCCCAGGCGAGGATGCCTGCC	709
Db	760	GCAACATCACAGACACCATGCTGTGCGCCAGTGTTCGAAAGAGGGCAAGGACTCCTGTC	819
QY	710	AGGTGATTCCTGGGGCCCCCTGTGTGTGGGGAGTCTTCAAGTCTGGTGTCTCTGGG	769
Db	820	AGGTGACTCTGGAGCCCCCTGTGTGTGCAACGATCTCTTCAAGGATCATCTCTCTGGG	879
QY	770	GGTCTGTGGGGCCCTGTGGAACAAGATGGCATCCCTGGAGTCTACCTATATTGCAAGT	829
Db	880	G---TCAGGACCCCATGTGCGTCAACAGAAAGCCTGGTGTCTATACAAAAGTCTGCAAT	936
QY	830	ATGTGACTGGATCCGGATGATCATGAGGAACAACTGACCTGTTCTCTCCACCTCCACC	889
Db	937	ACTTTAACTGGATCCACGAGGTTATGAGGAACAATTAGAGGGGACCTGCTTCCACCACC	996
QY	890	CCACCCCT	897
Db	997	CAACCCCT	1004

Search completed: June 20, 2004, 04:42:18  
Job time : 3358 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 02:20:26 ; Search time 99 Seconds  
(without alignments)  
6115.671 Million cell updates/sec

Title: US-10-015-385A-193

Perfect score: 1091

Sequence: 1 caagcaggtcatcccttgg.....aggggcaaaaaaaaaa 1091

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 27747546 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	506.4	46.4	840	4	US-09-244-111-5
2	493.2	45.2	619	4	US-09-280-116-139
3	219.4	20.1	1146	4	US-09-205-258-247
4	217.8	20.0	1166	3	US-08-944-483-7
5	217.6	19.9	1192	3	US-08-944-483-8
6	217.6	19.9	1314	3	US-09-025-059-2
7	216.8	19.9	1292	4	US-08-205-258-189
8	206.8	19.0	833	2	US-08-730-137-2
9	199.4	18.3	1052	4	US-09-386-642-10
10	191.6	17.6	1343	4	US-09-618-259-72
11	191.6	17.6	1360	4	US-09-618-259-6
12	190	17.4	994	3	US-09-008-271A-19
13	188.4	17.3	1049	4	US-09-386-642-9
14	184.8	16.9	944	3	US-09-070-526-1
15	173.8	15.9	1454	1	US-08-467-155A-2
16	173.8	15.9	1454	2	US-08-628-198-2
17	173.8	15.9	1454	3	US-09-201-038-2
18	173.8	15.9	1454	5	PCT-US96-07343-2
19	172.8	15.8	825	3	US-09-120-582-1
20	169.4	15.5	897	2	US-08-956-267A-1
21	162	14.8	732	1	US-08-361-395-2
22	146	13.4	1499	4	US-09-509-908-1
23	143	13.1	969	3	US-09-502-600-30
24	143	13.1	969	4	US-09-918-243-30
25	143	13.1	986	2	US-08-557-146-1
26	143	13.1	986	2	US-08-154-344-1
27	143	13.1	1089	3	US-08-930-188-1

C	28	143	13.1	1089	3	US-08-930-188-3	Sequence 3, Appli
C	29	143	13.1	1089	5	PCT-US96-04294-1	Sequence 1, Appli
C	30	143	13.1	1089	5	PCT-US96-04294-3	Sequence 3, Appli
	31	136	12.5	693	4	US-09-402-515A-17	Sequence 17, Appli
	32	135	12.4	1476	2	US-08-824-874-2	Sequence 2, Appli
	33	135	12.4	1476	3	US-09-210-084-2	Sequence 2, Appli
	34	135	12.4	1476	4	US-09-764-762-2	Sequence 2, Appli
	35	125.8	11.5	725	3	US-09-197-801-12	Sequence 12, Appli
	36	125.8	11.5	725	3	US-09-551-028-12	Sequence 12, Appli
	37	125.8	11.5	725	4	US-09-664-595A-12	Sequence 12, Appli
	38	122.2	11.2	792	4	US-09-244-111-3	Sequence 3, Appli
	39	118.6	10.9	1504	4	US-09-280-116-1	Sequence 1, Appli
	40	118	10.8	1341	4	US-08-983-075D-6	Sequence 6, Appli
	41	118	10.8	1358	4	US-08-983-075D-8	Sequence 8, Appli
	42	117.4	10.8	832	3	US-08-768-859A-5	Sequence 5, Appli
	43	117.4	10.8	832	3	US-08-768-859A-20	Sequence 20, Appli
	44	117.4	10.8	832	3	US-08-767-820A-5	Sequence 5, Appli
	45	117.4	10.8	832	3	US-08-767-820A-20	Sequence 20, Appli

#### ALIGNMENTS

#### RESULT 1

US-09-244-111-5  
; Sequence 5, Application US/09244111  
; Patent No. 6566498  
; GENERAL INFORMATION:  
; APPLICANT: Ni, et al.  
; TITLE OF INVENTION: Human Serine Protease and Serpin Polypeptides  
; FILE REFERENCE: PF391  
; CURRENT APPLICATION NUMBER: US/09/244,111  
; CURRENT FILING DATE: 1998-02-04  
; EARLIER APPLICATION NUMBER: 60/073,961  
; EARLIER FILING DATE: 1998-02-06  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 840  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (115)..(603)  
US-09-244-111-5

Query Match		46.4%	Score 506.4;	DB 4;	Length 840;
Best Local Similarity		97.9%	Pred. No. 3.4e-130;		
Matches 513;		Conservative	0;	Mismatches 11;	Indels 0; Gaps 0;
Qy	568	CACCCACGGAAACCCATTCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATGCTCCCAT	627		
Db	304	CAGCGGAGACCCATTCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATGCTCCCAT	363		
Qy	628	GCCACCTGCCATGGTGTGTATCCGGGAGAAATCACGAGCAACATGTTGTGCGAGCGGC	687		
Db	364	GCCACCTGCCATGGTGTGTATCCGGGAGAAATCACGAGCAACATGTTGTGCGAGCGGC	423		
Qy	688	GTCCCGGGGAGGATGCTCCAGGTGATCTTGGGGCCCCCTGGTGTGGGGGAGTC	747		
Db	424	GTCCCGGGGAGGATGCTCCAGGTGATCTTGGGGCCCCCTGGTGTGGGGGAGTC	483		
Qy	748	CTTCAAGTCTGTGTCTCTGGGGTCTGTGGGGCCCTGTGGCAAGATGGATCCCTCGGA	807		
Db	484	CTTCAAGTCTGTGTCTCTGGGGTCTGTGGGGCCCTGTGGCAAGATGGATCCCTCGGA	543		
Qy	808	GTCTACACCTATATTTGCAAGTATGTGAGATCGGATCGGATGATCATGAGGAAACAATGA	867		
Db	544	GTCTACACCTATATTTGCAAGTATGTGAGATCGGATCGGATGATCATGAGGAAACAATGA	603		
Qy	868	CTGTGTTTCTTCCACCTCCACCCACCCCTTAACCTTGGGTACCCCTCTGGCCCTCAGGC	927		
Db	604	CTGTGTTTCTTCCACCTCCACCCACCCCTTAACCTTGGGTACCCCTCTGGCCCTCAGGC	663		



EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 247  
LENGTH: 1146  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (20)  
OTHER INFORMATION: n equals a,t,g, or c  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (35)  
OTHER INFORMATION: n equals a,t,g, or c  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (36)  
OTHER INFORMATION: n equals a,t,g, or c  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (37)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-205-258-247

Query Match 20.1%; Score 219.4; DB 4; Length 1146;  
Best Local Similarity 58.5%; Fred. No. 7.1e-51;  
Matches 436; Conservative 3; Mismatches 294; Indels 12; Gaps 3;  
QY 168 GGCAGCCACACCGAAGATTTTCAATGGCACTGAGTGGGGCGTAACCTCACAGCCGTTGGCA 227  
DB 108 GGGGAGAGACCCAGAGATCATCAGGGTTGAGTGCAGCCCTCACTCCAGCCCTGGCA 167  
QY 228 GTGGGGCTTTTGGGGCAACAGCTGCGCTCGGGGGTGTCTTATTGACACACAGGTG 287  
DB 168 GGCAGCCCTTTCGAGAAGACCGGGCTACTCTGTGGGGCGACGCTCATCGCCCCAGATG 227  
QY 288 GGTCTTCACAGCGGCTCACTGACGGCAGCAGAGTACTGGGTGGCTGGGGGAACACAG 347  
DB 228 GCTCTTCAGAGCCGACCTGCTCAAGCCCGCTACATAGTTTCACTGGGGGAGACAA 287  
QY 348 CTTGAGCCAGCTCGATGGACCGAGAGATCCGGCAGCGGGCTTCTCTGTGACCCATCC 407  
DB 288 CTTCCAGAGAGAGGGCTGTGAGCAGACCCGACAGCCAGCTGAGTCTTCCCCACCC 347  
QY 408 CGGCTACCTGGAGCCCTCGAGAGGCCACAGAC-----GACTCGGCTGTGGGGCT 461  
DB 348 CGGCTTCAACACAGCCCTCCCAACAAAGACCCGCAATGACATCATGTCTGTGAAGAT 407  
QY 462 GCGCTGCGGCTCGGCTATACAGCAGCGTTCAACCCCTGCCCTGCCCAATGACTGTGC 521

DB 408 GGCATCGCCAGTCTCCATCACCTGGGCTGTGGACCCCTCACCTCTCTCACGCTGTGT 467  
QY 522 AACCGTGCACGAGTGCACGTCTCAGGTGGGGCATCAACACACCCACCGAACCC 581  
DB 468 CACTGTGGCACCAGCTGYCTCATTTCCGGCTGGGGCAGMACGTCCAGGCCCCAGTTACG 527  
QY 582 ATTCCCGGATCTGCTCCAGTGTCTCAACCTCTCCATCTCTCCATGCGCACCTGCCATGG 641  
DB 528 CTTGCTCACCTTGGGATGGCCCAATCACCATCTTGGACACAGAGTGTGAGAA 587  
QY 642 TGTGTATCCCGGAGAAATCAGAGCAACATGTGTGTGAGGGG---GGTCCCGGGGCA 698  
DB 588 CGCCTACCCCGGCAACATCACAGACACCATGTGTGTGCGAGCGTGCAGGAAGGGGGCAA 647  
QY 699 GATGCTGCGCAGGTGATCTGG 758  
DB 648 GGAATCTCTCCAGGTGACTTCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 707  
QY 759 GGTGTCTCTGGGGTCTGTGG 818  
DB 708 TATCTCTGGG---CCAGATCCGTGTGGGATCACCCGAAAGCCCTGGTGTCTACAGAA 764  
QY 819 TATTGCAAGTATGTGAGTCCGATCCGATGATCATGAGGAAACACTGACCTGTTTCTC 878  
DB 765 AGTCTGCAATATGTGAGTCCGATCCAGGAGACGATGAAGAAACAATTAGACTGGACCCAC 824  
QY 879 CACTCCACCCCGACCCCTTAATT 903  
DB 825 CCACACGCCCATCACCTCCATT 849

## RESULT 4

US-08-944-483-7  
Sequence 7, Application US/08944483  
Patent No. 6232456  
GENERAL INFORMATION:  
APPLICANT: COHEN, MAURICE  
APPLICANT: COLPITTS, TRACEY L.  
APPLICANT: FRIEDMAN, PAULA N.  
APPLICANT: GRANADOS, EDWARD N.  
APPLICANT: KLASS, MICHAEL R.  
APPLICANT: RUSSELL, JOHN C.  
APPLICANT: STEWART, KENT D.  
APPLICANT: STROUPE, STEVEN D.  
TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS  
TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES  
TITLE OF INVENTION: OF THE PROSTATE  
NUMBER OF SEQUENCES: 76  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/944,483  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Becker, Cheryl L.  
REGISTRATION NUMBER: 35,441  
REFERENCE/DOCKET NUMBER: 6183.US.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847/935-1729

TELEFAX: 847/938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1166 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-944-483-7

Query Match 20.0%; Score 217.6; DB 3; Length 1166;

Best Local Similarity 58.9%; Pred. No. 2e-50;

Matches 432; Conservative 1; Mismatches 289; Indels 12; Gaps 3;

```
QY 179 CGAGATTTCATGCGCTGAGTGTGGGCTTAATCTACAGCGGTGGAGTGGGGCTGT 238
Db 161 CCAGGATCATCAAGGGTTCAGTGCNAGCCCTCACTCCAGCCCTGGCAGGCGCTGT 220
QY 239 TTGAGGGACACGAGCTCGCTGCGGGGTGCTCTTATTGACACAGGTGGTCTTCACAG 298
Db 221 TCRAGAARACGGGTATCTGTGGGGGAGCTCATCGCCCCAGATGGCTCTCTGACAG 280
QY 299 CGGCTCACTGACGGGCGAGAGGTACTGGTGGCGCTGGGGGAAACACAGCTTCAGCCAGC 358
Db 281 CAGCCCACTGCTCAAGCCCGCTACATAGTTCACTGGGGCAGCACAACTCCAGAAGG 340
QY 359 TCGACTGACGAGCAGATCGGACAGCGGCTTCTGTGACCCATCCCGGCTACCTGG 418
Db 341 AGGAGGGGTGTGAGCAGACCGGACAGCCACTGTGCTCTCCCGACCCCGGCTTCAACA 400
QY 419 GAGCTCCAGCAGGACACAGACAC-----GACCTCGGCTGTGGGCTGGCGCTGCCG 472
Db 401 ACAGCTCCCCAACAAAGACACACCGCATGATCATCTGTTGAAGATGGCATCGCCAG 460
QY 473 TCCGGGTAACAGCAGCGTTCAACCCCTGCCCTGCCCCCAATGACTGTGCAACCGCTGGCA 532
Db 461 TCTCCATCACCTGGGCTGTGGGACCCCTCACCTCTCTCAGCTGTGTCTACTGTGGCA 520
QY 533 CCGAGTGCCAGTCTCAGGCTGGGGCATCACCAACCCACGGAACCCATTCCCGGATC 592
Db 521 CCAGTGTCTCATTTCCGGCTGGGGCAGACGTCAGCCCCCGAGTTAGCCTGCTCTACA 580
QY 593 TGCTCCAGTGCCTCAACCTCTCCATCGTCTCCATGCGACCTGCCATGTTGTGTATCCG 652
Db 581 CTTTGGATGGCCCAACATCACCATTTGAGCACCAGAGTGTGAGAACGCTTACCCTG 640
QY 653 GGAGNATCAGAGCAACATGTTGTGTGAGGG-----GGTCCCGGGGAGGATGCTGCC 709
Db 641 GCAACATCAGACACCACTATGTTGTGTGCGAGCGTGCAGGAAGGGGCAAGGACTCCTGCC 700
QY 710 AGGCTGATTTCTGGGGGCCCCCTGGTGTGTGGGGAGTCTCTCAAGGTCTGTGTCTGGG 769
Db 701 AGGCTGATTCGGGGGCCCCCTGGTGTGTGTAACTGCTCTCAAGGATTAATCTCTGG 760
QY 770 GGTCTGTGGGGCCCTGTGGCAAGATGGATCCCTGGAGTCTACACCTATATTTGCAAGT 829
Db 761 G-----CCAGGATCCGCTGTGCGATCACCCGAAAGCCCTGGTGTCTACACGAAAGTCTGCAAT 817
QY 830 ATGTGAGTGGATCCGATGATCATGAGGAACTGACCTGTTTCTTCCCTCCACCTCCACC 889
Db 818 ATGTGAGTGGATCCGAGGACGATGAAGAACTATGACTGGACCCCAACCAACCAAGCC 877
QY 890 CCACCCCTTAATTT 903
Db 878 CATCACCTCCATT 891
```

RESULT 5

US-08-944-483-8

Sequence 8, Application US/08944483

Patent No. 6232456

GENERAL INFORMATION:

APPLICANT: COHEN, MAURICE

```
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; TITLE OF INVENTION: OF THE PROSTATE
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183-US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-944-483-8
```

Query Match 19.9%; Score 217.6; DB 3; Length 1192;

Best Local Similarity 59.0%; Pred. No. 2.3e-50;

Matches 433; Conservative 0; Mismatches 289; Indels 12; Gaps 3;

```
QY 179 CGAGATTTCATGCGCTGAGTGTGGGCTTAATCTACAGCGGTGGAGTGGGGCTGT 238
Db 165 CCAGGATCATCAAGGGTTCAGTGCNAGCCCTCACTCCAGCCCTGGCAGGCGCTGT 224
QY 239 TTGAGGGACACGAGCTCGCTGCGGGGTGCTCTTATTGACACAGGTGGTCTTCACAG 298
Db 225 TCGAGAAGACGCGCTACTCTGTGGGGGAGCTCATCGCCCCAGATGGCTCTGACAG 284
QY 299 CGGCTCACTGACGGGCGAGAGGTACTGGGTGGCGCTGGGGGAAACACAGCTTCAGCCAGC 358
Db 285 CAGCCCACTGCTCAAGCCCGCTTACATAGTTTACCTGGGGGAGCAAACTCCAGAAGG 344
QY 359 TCGACTGACCGAGCAGATCGGACAGCGGCTTCTGTGACCCATCCCGGCTACCTGG 418
Db 345 AGGAGGGCTGTGAGCAGACCCGAGACCCACTGAGTCTTCCCCCAACCCCGGCTTCAACA 404
QY 419 GAGCTTCGACGAGCCACGAGCAC-----GACCTTCGGGCTGCTCGGCTGGCGCTGCCG 472
Db 405 ACAGCTCCCCAACAAAGACACCCGCAATGACATCATGCTGTTGAAGATGGCATCGCCAG 464
QY 473 TCCCGGTAACAGCAGCGTTCACCCCTGCCCTGCCCCCAATGACTGTGCAACCGCTGGCA 532
```

Db 465 TCTCCATCAGCTGGGCTGTGGAGCCCTCACCTCTCTCAAGCTGTGTCACTGCTGGCA 524  
QY 533 CCGAGTGCAGCTCTAGGCTGGGGCATCACCAACCCACCCAGCAACCCATTCCCGGATC 592  
Db 525 CCAGCTGCCTCATTTCCGGCTGGGGCAGCAGCTCCAGCCCCCAGTTACGGCTGCTCACA 584  
QY 593 TGTCCAGTGCCTCAACTCTCCATCTCCATCTCCATGCTCCATGCTCCATGCTGTATCCCG 652  
Db 585 CTTTGGCATGGCCAAACATCAACCAATCATGTAGCAGCAGAAAGTGTGAGAACGCTACCCG 644  
QY 653 GGAATATCAGAGCAACATGTGTGTGAGGCG---CGGTCCCGGGGAGGATCCCTGCC 709  
Db 645 GCAACATCAGACAGACCATGTGTGTGCGAGCGTGCAGAAAGGGGCAAGGACTCCTGCC 704  
QY 710 AGGTGATCTGGGGGCCCCCTGTGTGTGGGGAGTCTTCAAGGTCTGTGTCTGGG 769  
Db 705 AGGTGATCTGGGGGCCCCCTGTGTGTGTAAACAGTCTTCAAGGCAATATCTCTGGG 764  
QY 770 GGTCTGTGGGGCCCTGTGCAAGATGGCATCCCTGGAGTCTACACCTATATTTGCAAGT 829  
Db 765 G---CCAGATCCGTGTGCGATCACCCGAAAGCCTGTGTCTACAGAAAGTCTGCAAT 821  
QY 830 ATGTGATGTGATCCGAGATCATGAGAAACATGTAACCTGTTTCTTCCACTCCAGCC 889  
Db 822 ATGTGATGTGATCCGAGAGCATGAAGAAACATTAAGTGTGACCCACCCACACAGCC 881  
QY 890 CCACCCCTTAATT 903  
Db 882 CATCACCTCCATT 895

## RESULT 6

US-09-025-059-2  
; Sequence 2, Application US/09025059  
; Patent No. 6075136  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; TITLE OF INVENTION: PROSTATE-ASSOCIATED SERINE PROTEASE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/025,059  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0481 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1314 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single

; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: LUNGUTU10  
; CLONE: 2723646  
US-09-025-059-2  
  
Query Match 19.9%; Score 217.6; DB 3; Length 1314;  
Best Local Similarity 59.0%; Pred. No. 2.4e-50;  
Matches 433; Conservative 0; Mismatches 289; Indels 12; Gaps 3;  
  
QY 179 CGAAGATTTCAATGGCACTAGTGTGGGCGTAACTCAGACCGCTGCGAGGTGGGGTGT 238  
Db 282 CCAGGATCATCAAGGGGTTCCAGTGAAGGCTCACTCCAGCCCTGGCAGCAGCCCTGT 341  
QY 239 TTGAGGGGCAACAGCCTCGCTGCGGGGCTGTCTTATTGACCAAGTGGTGTCTCTCAG 298  
Db 342 TCGAAGAAGACCGGCTACTCTGTGGGGGAGCTCATCGCCCCAGATGGTCTCTGACAG 401  
QY 299 CGGCTCACTGAGCGGAGCAGGTACTGGGTGCGCTTGGGGAAACAGAGCTTCAGCCAGC 358  
Db 402 CAGCCCACTGCCTCAAGCCCGCTACATAGTTCACTGGGGGAGCAACAACCTCCAGAAG 461  
QY 359 TCGAGTGCAGCGACAGATCGGSCAGAGCGGCTTCTGTGACCCATCCCGGCTACCTGG 418  
Db 462 AGGAGGGCTGTGACAGACCCGGAAGCCACTGAGTCTTCCCGGAGCCCGGCTTCAACA 521  
QY 419 GAGCCTCGAGGAGCAGAGCAC-----GACCTCGGGCTGTGGGCTGCGCTGCCCG 472  
Db 522 ACAGCTCCCCAACAGAGACACCGCAATGATCATGCTGTGAAGATGGCATCGCCAG 581  
QY 473 TCCGCGTAACAGCAGCGGTTCAACCCCTGCGCCCTGCCAAATGACTGTGCAACCGGTGGCA 532  
Db 582 TCTCATCACTGGGCTGTGGAGCCCTCAACCTCTCCTCAGCTGTGTCACTGTGGCA 641  
QY 533 CCGAGTGCAGCTCTCAGGCTGGGGCATCACCAACCCAGCAAGCCATTCGCCGATC 592  
Db 642 CCAGCTGCCTCATTTCCGGGCTGGGGCAGCAGTCCAGCCCCAGTTACGCTTGCCTCACA 701  
QY 593 TGTCCAGTGTCTCAACCTCTCCATCGTCTCCATGCCCATGCCATGCTGTGTATCCCG 652  
Db 702 CTTTGGATGGCCCAACATCACTACCATCATGTAGCACCAGCAAGTGTGAGAACGCTACCCG 761  
QY 653 GGAGATCAGAGCAACATGTGTGTGAGGCG---GGTCCCGGGGAGGATGCTGCC 709  
Db 762 GCAACATCAGACAGACCATGTGTGTGCGAGCGTGCAGAAAGGGGCAAGGACTCCTGCC 821  
QY 710 AGGTGATCTTGGGGGCCCCCTGTGTGTGGGGAGTCTTCAAGGTCTGTGTCTCTGG 769  
Db 822 AGGTGATCTTGGGGGCCCCCTGTGTGTGTAAACAGTCTTCAAGGCAATATCTCTGG 881  
QY 770 GGTCTGTGGGGCCCTGTGGAACAAGATGGCATCCCTGGAGTCTACACCTATATTTGCAAGT 829  
Db 882 G---CCAGATCCGTGTGGATCACCCGAAAGCCTGGTGTCTACAGAAAGTGTGCAAT 938  
QY 830 ATGTGACTGTATCCGAGATGATCATGAGAAACATGACTGTGTCTTCCCTCCACTCCAGCC 889  
Db 939 ATGTGACTGTATCCGAGAGAGATGAAGAAACATTAAGTGTGACCCACCCACACAGCC 998  
QY 890 CCACCCCTTAATT 903  
Db 999 CATCACCTCCATT 1012

## RESULT 7

US-09-205-258-189  
; Sequence 189, Application US/09205258  
; Patent No. 6525174  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 207 Human Secreted Proteins  
; FILE REFERENCE: P2007P1  
; CURRENT APPLICATION NUMBER: US/09/205,258  
; CURRENT FILING DATE: 1998-12-04

EARLIER APPLICATION NUMBER: PCT/US98/11422  
EARLIER FILING DATE: 1998-06-04  
EARLIER APPLICATION NUMBER: 60/048,885  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,375  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,881  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,880  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,896  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,020  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,876  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,895  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,884  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,894  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,971  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,882  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,899  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,893  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,900  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,901  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,892  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,915  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,019  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,972  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,916  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,373  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,875  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,374  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,917  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,949  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878

EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 189  
LENGTH: 1292  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-205-258-189

Query Match 19.9%; Score 216.8; DB 4; Length 1292;  
Best Local Similarity 58.7%; Pred. No. 3.9e-50;  
Matches 431; Conservative 2; Mismatches 289; Indels 12; Gaps 3;

QY 179 CGAAGATTTCATATGCACTGAGTGTGGGGGTAACTCAGACGGCTGGCAGGTGGGCTGT 238  
DB 265 CCAGGATCATCAGGGGTTCGAGTCAAGCTCCTCAGCCCTGGCAGGCGAGCCTGT 324  
QY 239 TTGAGGGCCACGACCTGCGCTGCGGGGTGTCTTATTGACCACAGGTGGGTCTCACAG 298  
DB 325 TCGAGAAGACGGGGTACTCTGTGGGCGAGCTCATCGCCCCACAGATGGCTCTGACAG 384  
QY 299 CGGCTCACTGACGGCAGAGGTACTGGGTGGCTTGGGGGAAACACAGCTCAGCCAGC 358  
DB 385 CAGCCCACTGCTCAAGCCCGCTACATAGTTTCACTGGGGCAGCACACTCCAGAAAG 444  
QY 359 TCGACTGGACCGAGCAGATCCGGCAGACGGGCTTCTGTGACCCCATCCGGGTACTGG 418  
DB 445 AGAGGGCTGTGAGCAGACCCGAGACAGCCACTGAGTCTTCCCCCAGCCCGGTCAACA 504  
QY 419 GAGCCTCGAGAGCCACGAGCAC-----GACCTCGGGTGTGCGGCTGCGCTGCCCG 472  
DB 505 ACAGCCTCCCAACAAGACCCGCAATGACATCATGTGTGTAAGATGGCATCGCCAG 564  
QY 473 TCCGGTAAACAGCAGGTTCAACCCCTGCGCTGCGCATGACTGTGCAACCGCTGGCA 532  
DB 565 TCTCCATCACTGGGCTGTGCGACCCCTCACCTCTCTCCTCAGCTGTGTCACTGTGGCA 624  
QY 533 CCGAGTGCCACGCTCTCAGGCTGGGGCATCACCAACACCCAGCAGCTCCAGCCCTGATC 592  
DB 525 CCAGCTGYCTATTTCGGCTGGGGCAGCAGCTCCAGCCCTGATGAGAACGCTACCCCG 684  
QY 593 TGCTCCAGTGCCTCAACCTCTCCATGCTTCCCATGCGCAGCTGCGATGGTGTATCCCG 652  
DB 685 CCTTGGGATGCGGCAACATCACCATCATGAGCAGCAGAGTGTGAGAACGCTACCCCG 744  
QY 653 GGAGAAATCAGAGCAGCATGTGTGCGAGCG---CGTCCCGGGGCGAGATGCGCTGCC 709  
DB 745 GCAACATCAGACACCATGGTGTGCGCAGCTGCGAGGAGGGGGGCAAGATCTCTGCC 804  
QY 710 AGGCTGATTCTGGGGGCCCCCTGGTGTGGGGAGTCTTCAAGTCTGGTGTCTGGG 769  
DB 805 AGGCTGACTCGGGGGCCCTCTGCTGTGTAACCACTGCTTCAAGGCAATATCTCTGGG 864  
QY 770 GGTCTGTGGGGCCCTGTGAGCAAGATGGCATCCCTGGAGTCTACCTATATTGCAAGT 829  
DB 965 G---CCAGGATCCGTGTGCGATCACCCGAAAGCTGTGTCTACAGAAAGTCTGCAAT 921  
QY 830 ATGTGACTGGATCCGGATGATCATGAGGAACAACTGACCTGTTTCTCCACCTCCACCC 889  
DB 922 ATGTGACTGGATCCAGGAGACCATGAGAACAAATTAGTCTGAGCCACCACCAAGSC 981  
QY 890 CCACCCCTTAACCTT 903  
DB 982 CATCACCTCCATT 995

RESULT 8



US-08-790-137-2  
; Sequence 2, Application US/08790137  
; Patent No. 5840871  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Goli, Surya K.  
; TITLE OF INVENTION: A NOVEL PROSTATE-ASSOCIATED  
; TITLE OF INVENTION: KALLIKREIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/790,137  
; FILING DATE: Filed Herewith  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0195 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-790-137-2

Query Match 19.0%; Score 206.8; DB 2; Length 833;  
Best Local Similarity 56.6%; Pred. No. 1.8e-47;  
Matches 442; Conservative 0; Mismatches 327; Indels 12; Gaps 3;

QY 120 CATGGGGCTCAGATCTTTTGGCTCTCTGTGTGTCTTGGGCTCAGCCAGCAGCCACACC 179  
DB 35 CTGGTTCTGTGCTCGCCCTGTCCCTGGGGGGAGTGTGTGTGCTGCGCCCGGATTGATC 94  
QY 180 GAAGATTTTCAATGGCACTGAGTGTGGGGGTAACTCAGAGCGGTGGAGTGGGGCTGT 239  
DB 95 CCGATTTGGGAGGCTGGAGTGTGAGAGCATTTCCAGCCCTGCGAGCGGCTCTGTA 154  
QY 240 TGAGGGCACCAGCTCGCTGCGGGGGTGTCTTATTTGACCAAGTGGGTCTCTCAGC 299  
DB 155 CCAGAAGACGCGCTACTCTGTGGGCGCAGCAGTCAATGCCCCAGATGTTCTGACAGC 214  
QY 300 GGCTCTACTGACGGGAGAGGAGTGTGGGTGCGCTGGGGGAAACAGAGCTTACGAGCT 359  
DB 215 AGCCCACTGCTTAAAGCCCGCTACATAGTTTCACTGGGGCAGCAGCAACCTCAGAAGGA 274  
QY 360 CGACTGAGCCAGCAGATCCGGCAGAGCGGCTTCTGTGACCCATCCCGGCTACCTGGG 419  
DB 275 GGAGGGCTGTGAGCAGACCCGAGCAGCCTAGTCTTCTCCCAACCCCGGCTTCAACAA 334  
QY 420 AGCCTCGACAGCCACAGCAGC-----GACCTCCGGCTCTCGGCTCGCCCTGCCCGT 473  
DB 335 CAGCCTCCCCAACAAAGACACCGCAATGACATCATGCTGGTGAAGATGGCATCGCCAGT 394  
QY 474 CCGGTAAACAGCAGGTTTCAACCCCTGCCCTTCCCAATGATGTGACACCGCTGGCAC 533

DB 395 CTCCATCACCTGGGCTGTGCGACCCCTCACCTCTCTCTCAAGCTGTGTCACTGTGGCAC 454  
QY 534 CGAGTGCCACGCTCTCAGGCTGGGGCATCAACACCCAGCGAACCCATTCGCGATCT 593  
DB 455 CAGTGGCTCATTTCCGGCTGGGGCAGCAGCTCAGCCCCCAGATTACGCTGCCTCACAC 514  
QY 594 GCTCCAGTGCCTCAACCTCTCCATCTCCCATGCCACCTGCCATGGTGTGTATTCGCG 653  
DB 515 CTTGGGATGCGCCACATCACCATTGAGCACCAGAGTGTGAGAACGCTACCCCG 574  
QY 654 GAGATCAGAGCAACATGGTGTGTGAGGGG---GCGTCCCGGGGAGGATGCTGCA 710  
DB 575 CAACATCAGACACCATGGTGTGTGCCAGGCTGAGGAAGGGGCAAGGACTCTCGCA 634  
QY 711 GGGTGAATTCTGGGGGCCCCCTGGTGTGTGGGGAGTCTTCAAGGTCTGGTCTCTGGG 770  
DB 635 GGGTGAATTCTGGGGGCCCCCTGGTGTGTAAACAGTCTCTTCAAGGATTTCTCTGGG 694  
QY 771 GTCTGTGGGGCCCCCTGTGGACAGATGGCATCCCTGGAGTCTACACTATATTGCAAGTA 830  
DB 695 ---CCAGGATCCGTGTGCGATCACCAGAAAGCTGTGTCTACAGAAAGTCTGCAATA 751  
QY 831 TGTGACTGGATCCGGATGATCATGAGGAAACACTGACTGTCTTCTCAGCTTCCACCTC 890  
DB 752 TGTGACTGGATCCAGGAGACGATGAGAACAAATTAGACTGGACNTACCTCCGAANCC 811  
QY 891 C 891  
DB 812 C 812

## RESULT 9

US-09-386-642-10  
; Sequence 10, Application US/09386642  
; Patent No. 6420157  
; GENERAL INFORMATION:  
; APPLICANT: Darrow, Andrew  
; APPLICANT: Qi, Jensen  
; APPLICANT: Andrade-Gordon, Patricia  
; TITLE OF INVENTION: Zymogen Activation System  
; FILE REFERENCE: ORT-1028  
; CURRENT APPLICATION NUMBER: US/09/386,642  
; CURRENT FILING DATE: 1999-08-31  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 1052  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene  
; OTHER INFORMATION: with homo sapien serine protease catalytic domain  
US-09-386-642-10

Query Match 18.3%; Score 199.4; DB 4; Length 1052;  
Best Local Similarity 59.6%; Pred. No. 2.3e-45;  
Matches 395; Conservative 0; Mismatches 256; Indels 12; Gaps 3;

QY 212 ACTCAGCGCTGGCAGGTGGGGCTGTTTGGGGCACCAGCCTGCGCTGCGGGGTGTCC 271  
DB 197 ACTCCAGCCCTGGCAGGAGCCCTGTTGGAAGACGCGGCTACTCTGTGGGGCAGCG 256  
QY 272 TTATTGACCACAGGTGGGTCTCTCACAGGGCTCACTTGACGGCGCAGCAGGTACTGGGTG 331  
DB 257 TCATGCCCCAGATGGTCTCTGACAGAGCCCATCTGCTCAAGCCCCGTACATAGTTC 316  
QY 332 GCGTGGGGAAACAGCCTCAGCCAGTCACTGACGAGCAGATCCGGCAGCAGCGGT 391  
DB 317 ACCTGGGGCAGCACACCTCCAGAAAGGAGGGGTGTGAGCAGACCCCGCAGCCACTG 376  
QY 392 TCTCTGTGACCATCCCGCTACCTGGGAGCTTCAGCAGAGCCACGAGCAC-----GACC 445  
DB 377 AGTCTTCTCCCAACCCCGGCTTCAACACAGCGCTTCCCAACAAAGACCAACCGCAATGACA 436

QY	446	TCCGGCTGCTCGGGCTGGCCCTGCCGCTCCGGTAAACAGCAGCGTTCAACCCCTGCCCC	505
Db	437	TCATGCTGGTGAAGATGGCATCGCCAGTCTCCATCACCTGGGCTGTGGAGCCCTCACCC	496
QY	506	TGCCCAATGACTGTGTCAACCGCTGSCACCGAGTGCCAGTCTCAGGCTGGGGCATCACCA	565
Db	497	TCTCCTCAGCGTGTGTCACTGCTGSCACCAAGTGCCTCATTTCCGGCTGGGCGACGAGT	556
QY	566	ACCAACCACGGAAACCATTCGGGATCTGCTCCAGTGCCTCAACCTCTCCATTCGTCTCCC	625
Db	557	CCAGCCCCCAGTTACGCGTGCCTCACACTTTCGATGCGCCAAACATCACCATCATTTGAGC	616
QY	626	ATGCACCTGCGCATGGTGTGTATCCCGGGAGAATCACAGACAAACATGTGTGTGCAGCG	685
Db	617	ACCAGAAGTGTGAGAACGCTACCCCGGCAACATCACAGACACCATGTGTGTGCCACGG	676
QY	686	---CCGTCCCGGGCAGGATCGCTGCCAGGCTGATCTGGGGGGCCCCCTGGTGTGTGGGG	742
Db	677	TGCAGGAAGGGGGCAAGACTCCTGCCAGGCTGACTCCGGGGGGCCCTCTGGTCTGTAAAC	736
QY	743	GAGTCCCTTCAAGGCTGTGTGTCTCGGGGAGTGTGTGGGGCCCTGTGGACAAGATGGCATCC	802
Db	737	AGTCTCTCAAGGCATTATCTCTGGGG---CCAGGATCCGTGTGTGCGATCACCCGAAAGC	793
QY	803	CTGAGTCTACACCTATATTTTGCAGATATGTGGACTGATCCGGATGATCATAGGAGACA	862
Db	794	CTGTGTCTACACGAAAGTCTGCAAATATGTGGACTGGATCCAGGAGACGATGAAGAACA	853
QY	863	ACT	865
Db	854	ATT	856

```

RESULT 10
US-09-618-259-72/c
; Sequence 72, Application US/09618259
; Patent No. 6642013
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: No. 6642013el Extracellular Serine Protease
; FILE REFERENCE: D6020C1F2
; CURRENT APPLICATION NUMBER: US/09/618,259
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 09/127,444
; PRIOR FILING DATE: 1998-08-21
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 72
; LENGTH: 1343
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Anti-sense sequence of TADG-14
US-09-618-259-72

```

	Query Match	17.6%	Score 191.6;	DB 4;	Length 1343;
	Best Local Similarity	56.3%;	Prod. No. 3.6e-43;		
	Matches 400;	Conservative 0;	Mismatches 304;	Indels 6;	Gaps 2
QY	157	GGGCTCAGCCAGGCGACACACCGAGATTTTCATGSCACTGAGTGTGGCGTAAC	216		
DB	759	GGACATCCAGGGACACAGGAGGACAGGTGCTGGGGGGTCATGAGTGCCAAACCCCAITTCG	700		
QY	217	CAGCCGTGGCAGGTGGGCTGTTTGAGGCGACACGCTCGCGTCGGGGGGTGTCCCTATT	276		
DB	699	CAGCCTTGGCAGGGGGCCCTGTTCAGGGCCAGCAACTACTCTGTGGCGGTGCTCCTGTA	640		
QY	277	GACCACAGGTGGTCTTCACAGCGCTCACTGCAGCGGCAGCAGTACTGGGTGGCGCTG	336		
DB	639	GGTGGCACTGGTCTTTACAGCTGCCCACTGTAAAAACCGAATAACACAGTAGCGCCTG	580		
QY	337	GGGAAACACAGCCTTAGCCAGCTCGACTGGACCGAGCAGATCCGGCAGCAGGGGGCTTCTCT	396		

579	Db	GGAGACCA	CGCCTACAGATAAAGATGGCCAGACAGAAATACTGTGGTTCAGTCC	520
397	Qy	GTGACCCATCCGGCT---	ACCTGGAGCCTCGACGAGCCACGAGCAGCAGCCTCCGGCTG	453
519	Db	ATCCCAACCCCTGCTACAACAGCAGCGATGTGGAGACCACAACTGATCTGTGTCTT	460	
454	Qy	CTCGGCTGCGCTGCGCGTCCGGGTAAACCAAGCAGCGTTCAACCCCTGCCCTGCCCAAT	513	
459	Db	CTTCAACTGCGTGACACAGGCATCCCTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGAT	400	
514	Qy	GACTGTGCAACGGCTGGCACCGAGTGCACGTCTCAGGCTGGGGCATCACCAACCA	573	
399	Db	CATTGCACCCAGCCTGGCCAGAAAGTGCACCGTCTCAGGCTGGGGCATGTCAACAGTCC	340	
574	Qy	CGGAACCCATTCCGGGATCTCTCCAGTGCCTCAACCTTCCATCGTCTCCCATGCCAC	633	
339	Db	CGAGAGATTTTCTGTGACACTCTCACTGTGCAGAGTAAATCTTTCCCAAGAGAAG	280	
634	Qy	TGCCATGTGTGTATCCGGGAGAATCAGCAGCAACATGGTGTGTGACGGCGCTCCGG	693	
279	Db	TGTGAGGATGCTTACCCGGGCGAGATCACAGATGGCATGCTGTGTCAGGCACAGCAAA	220	
694	Qy	GGCAGGATGCTGCCAGGTGATCTTGGGGCCCCCTGTGTGTGGGGAGTCTCTCAA	753	
219	Db	GGGGGTGACAGTGGCCAGGGCGATCTTGGAGGCCCCCTGTGTGTGTGTGATGGTGCAC	160	
754	Qy	GGTCTGGTGTCTCTGGGGGTCTGTGGGGCCCTGTGGACAAGATGGCATCCCTCGAGTCTAC	813	
159	Db	GGCATCACTCTGGGGTCTC---AGA	103	
814	Qy	ACCTATATTTGCAAGTATGTGCACTGGATCCGGATGATCATGAGGAACAA	863	
102	Db	ACCAACATCTCGCGTACTCGATCTGGATCGGATCAAGAAGATCATAGCAGCA	53	

```

RESULT 11
US-09-618-259-6
; Sequence 6, Application US/09618259
; Patent No. 6642013
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: No. 6642013el Extracellular Serine Protease
; FILE REFERENCE: D6020CIP2
; CURRENT APPLICATION NUMBER: US/09/618,259
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 09/127,444
; PRIOR FILING DATE: 1998-08-21
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 6
; LENGTH: 1360
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence encoding Tumor Antigen Derived
; OTHER INFORMATION: Gene-14 (TAGD-14) protein; nt 1344-1360 NCBI accession
; OTHER INFORMATION: #AA343629
; US-09-618-259-6

```

	Query Match	17.6%;	Score 191.6;	DB 4;	Length 1360;
	Best Local Similarity	56.3%;	Pred. No. 3.6e-43;		
	Matches 400;	Conservative	0;	Mismatches 304;	Indels 6;
				Gaps	2
QY	157	GGGCTCAGCCAGGCGCACACCGAAGATTTTCATGGCACTGAGTGTGGCGGTAACTCA	216		
Db	585	GGACACTCCAGGCGCAGGAGGACAGGTGCTGGGGGTCATGAGTGCCAAACCCCAATTCG	644		
QY	217	CAGCCGTGCGAGGTGGGCTCTTTTGAGGGCACACAGCTGGCTCGCGGGGTGTCCTATT	276		
Db	645	CAGCCTTGCGAGGGGGCCCTGTTCCAGGGCCAGCACTACTGTGGCGGTGTCCTTGTA	704		
QY	277	GACCACAGGTGGGTCTTCACACGGGCTCACTGCAGCGGCAGCAGGTACTGGGTGGCCCTTG	336		

Db 705 GGTGGCAACTGGTCTTACAGTGGCTGAAACCGAATACACACTAGCGCTG 764  
QY 337 GGGGAACACAGCGCTCAGCCAGCTGAGTGGACCGAGAGATCGGCACAGCGCTTCTCT 396  
Db 765 GGAGACACACAGCGCTACAGAAATAGATGGCCAGAGCAAGAAATCCTGTGGTTAGTCC 824  
QY 397 GTGACCCATCCCGGT--ACCTGGGAGCTCGACGAGCCAGAGCAACGACCTCGGCTG 453  
Db 825 ATCCACACCCCTGCTACACAGCAGCGATGGAGGACCAACCATATCTGATCTT 884  
QY 454 CTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 513  
Db 885 CTTCAACTGGTGAACGAGCATCCCTGGGCTCCAAAGTGAAGCCCATCAGCCTGGCAGAT 944  
QY 514 GACTGTGCAACCGCTGCGACCGAGTGGCAGCTCTCAGGCTGGGGCATCACCACACCA 573  
Db 945 CATTCACCCAGCGCTGCGCAGAGAGTGCACCGCTCTCAGGCTGGGGCATCTGCCAGTCCC 1004  
QY 574 CGGAACCCATTCGCGGATCTGCTCCAGTGCCTCAACCTCTCCATGCTCCCATGCCACC 633  
Db 1005 CGAGAGAAATTTCTGACACTCTCAACTGTGCAAGTAAATCTTTCCCGCAGAGAG 1064  
QY 634 TGCCATGGTGTATCCCGGAGAAATCACAGCAACATGTTGTGTCAGCGCGCTGCCG 693  
Db 1065 TGTGAGGATGCTTACCGGGGCGAGATCACAGATGGCATGCTGTGTCAGCGCAGCAAA 1124  
QY 694 GGGCAGGATGCTCCAGAGGTGATTTGGGGGCCCCCTGTTGTGTTGGGGAGTCTTCAA 753  
Db 1125 GGGGCTGACACGTCGCGCAGGCGGATCTGAGAGGCCCCCTGTTGTGTTGATGGTGCAC 1184  
QY 754 GGTCTGGTGTCTCGGGGCTGTGGGGCCCTGTGGCAAGATGGCATCCCTGGAGTCTAC 813  
Db 1185 GGCATCACATCTCGGGGCTC--AGACCCCTGTGGAGTCCGACAAACTTGGCGTCTAT 1241  
QY 814 ACCTATTTGCAAGTATGTGAGTCCGATCGATCATGAGGAACAA 863  
Db 1242 ACCAACATCTGCGCTACCTGGACTGGATCAAGAGATCATAGGCAGCAA 1291

## RESULT 12

US-09-008-271A-19  
Sequence 19, Application US/09008271A  
Patent No. 6203979

## GENERAL INFORMATION:

APPLICANT: Bandman, Olga

Hillman, Jennifer L.

Yue, Henry

Guegler, Karl J.

Corley, Neil C.

Tang, Tom Y.

Shah, Purvi

TITLE OF INVENTION: HUMAN PROTEASE MOLECULES

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Dr.

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/008,271A

FILING DATE: 16-Jan-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Mohan-Peterson, Sheela  
REGISTRATION NUMBER: 41,201  
REFERENCE/DOCKET NUMBER: PF-0458 US

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 994 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: COLNOT27

CLONE: 1798496

SEQUENCE DESCRIPTION: SEQ ID NO: 19:

US-09-008-271A-19

Query Match 17.4%; Score 190; DB 3; Length 994;

Best Local Similarity 56.2%; Pred. No. 8 7e-43;

Matches 399; Conservative 0; Mismatches 305; Indels 6; Gaps 2;

QY 157 GGGCTCAGCCAGCGCAGCCACACCGAAGATTTTCAATGGCACTGAGTGTGGCGTAAC 216  
Db 224 GGACACTCCAGGGCACAGGAGCAAGGTGCTGGGGGGTCAATGAGTGCACACCCATTCG 283  
QY 217 CAGCGTGGCAGGTGGGCTGTTTGGGGCACACGCTCGGTCGGGGGTGCTCTATT 276  
Db 284 CAGCCTTGGCAGCGGCTTGTCCAGGGCCAGCACTACTCTGTGGGGTGTCTTTGTA 343  
QY 277 GACCAAGTGGTCTCTCACAGCGCTCACTCAGCGGCGAGGATCTGCGTGGCGCTG 336  
Db 344 GGTGGCACTGGTCTTACAGCTGCCACTGTAAACCCGAAATACACAGTACGCGCTG 403  
QY 337 GGGGAACACAGCCTCAGCCAGCTCGACTGGAACGAGCAGATCCGGCAGCGCTTCTCT 396  
Db 404 GGAGACCAAGCCTACAGAAATAAGATGGCCAGAGCAAGAAATACCTGTGTTAGTCC 463  
QY 397 GTGACCCATCCCGCT--ACCTGGGAGCTCGACGAGCGCTGAGGAGCGATGATGATGCT 453  
Db 464 ATCCACACCCCTGTCTACAAAGCAGCGATGAGGAGCGATGATGATGATGATGCT 523  
QY 454 CTGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 513  
Db 524 CTTCAACTGGTGACCGAGCATCCCTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGAT 583  
QY 514 GACTGTGCAACCGCTGGCACCGAGTGCACGCTCTCAGGCTGGGGCATCACCACCA 573  
Db 584 CATTCACCCAGCCTGGCCAGAAAGTGCACCGCTCTCAGGCTGGGGCATCTGTACCACTCC 643  
QY 574 CGGAACCCATTCGCGATCTGCTCCAGTGCCTCAACCTCTCCATGCTCTCCATGCCACC 633  
Db 644 CGAGAGATTTTCTGACACTCTCAACTGTGAGAGTAAATCTTTCCCGAAGAGAG 703  
QY 634 TGCCATGGTGTATCCCGGAGAAATCACAGCAACATGTTGTGTCAGCGCGCTCGCG 693  
Db 704 TGTGAGATGCTTACCCGGGCGAGATCAAGATGGCATGCTGTGTGTCAGCGCAGCAAA 763  
QY 694 GGGCAGGATGCTGCCAGGGTATTTCTGGGGCCCCCTGTTGTGTTGGGGAGTCTCTTCAA 753  
Db 764 GGGGCTGACACGTGCCAGGGCGATTTCTGGAGGCCCTGTTGTGTTGATGTTGATGCT 823  
QY 754 GGTCTGGTCTCTGGGGTCTGTGGGGCCCTGTGGACAGATGGGATCCTCGAGTCTAC 813  
Db 824 GGCATCACATCTTGGGGCTC--AGACCCCTGTGGAGTCCGACAAACTTGGCGTCTAT 880  
QY 814 ACCTATTTGCAAGTATGTGAGTCCGATGATCCGGATCATGAGGAACAA 863  
Db 881 ACCAACATCTGCGCTACCTGGACTGGATCAAGAGATCATAGGCAGCAA 930

## RESULT 13

US-09-386-642-9

; Sequence 9, Application US/09386642

; Patent No. 6420157

; GENERAL INFORMATION:

; APPLICANT: Darrow, Andrew

; APPLICANT: Qi, Jensen

; APPLICANT: Andrade-Gordon, Patricia

; TITLE OF INVENTION: Zymogen Activation System

; FILE REFERENCE: ORT-1028

; CURRENT APPLICATION NUMBER: US/09/386,642

; CURRENT FILING DATE: 1999-08-31

; NUMBER OF SEQ ID NOS: 60

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 9

; LENGTH: 1049

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene

; OTHER INFORMATION: with homo sapien serine protease catalytic domain

US-09-386-642-9

Query Match 17.3%; Score 188.4; DB 4; Length 1049;

Best Local Similarity 57.4%; Pred. No. 2.5e-42;

Matches 380; Conservative 0; Mismatches 276; Indels 6; Gaps 2;

QY 212 ACTCAGCGGTGGGAGGTGGGGCTGTTTGGAGGACACAGCGCTGGCGTGGGGGGTGTCC 271

DB 197 ATTGGAGCCTTGGGAGGGGGCTGTTTCCAGGGCCAGCACTCTCTGTGGGGTGTCC 256

QY 272 TTATTGACCACAGGTGGGTCTTACAGCGGCTCACTGACGGCGCAGCGTACTGGGTGC 331

DB 257 TTGTAGTGGCACTGGGTCTTACAGCTGCCCTGTAAAAACCGAATAACAGTAC 316

QY 332 GCGTGGGGAACACAGCCTCAGCAGCTCGACTGACCGCAGCAGATCCGGCAGCGGCT 391

DB 317 GCGTGGGACACAGCCTACAGATAAAGATGGCCAGCAGAAATACCTGTGGTTTC 376

QY 392 TCTGTGTACCATCCGGCT---ACCTGGAGCCTCGACGAGCCACGACGACGACCTCC 448

DB 377 AGTCCATCCACACCCCTGTACACAGCAGCGATGTGGAGGCCACACACCATGATCTGA 436

QY 449 GGTGTGGGTGGTGGCTCCCGTCCCGGTAAACAGCAGCGGTTCACCCCTGGCCCTGC 508

DB 437 TGCTTCTTCAACTGCGTGACCGGATCCCTGGGTCCAAAGTGAAGCCCATCAGCCTGG 496

QY 509 CCAATGACTGTGCAACCGCTGGCAGCGAGTCCAGCTTCAGGCTGGGCGATCACCAACC 568

DB 497 CAGATCATTTGACCCAGCTGGCCAGAGTGACCGTCTCAGGCTGGGGCACTGTACCA 556

QY 569 ACCACGGAACCAATCCCGGATCTCTCCAGTGCCTCAACCTCTCCATCGTCTCCCATG 628

DB 557 GTCCCGGAGAGATTTTCTTGACACTCTCAACTGTGCAAGTAAATAATCTTTCCCCAGA 616

QY 629 GCACCTGCATGTGTGTATCCCGGAGATCACAGCAGCACTGTGTGTGAGGGGGGG 688

DB 617 AGAAGTGTGAGGATGTATTACCGGGGCGCATCACAGATGGCATGGTGTGTGATGGTGCAC 676

QY 689 TCCCGGGGAGGATGTCTCCAGGTGTATCTGGGGGCCCCCTGGTGTGTGGGGAGTCC 748

DB 677 GCMAAGGGGTGACAGTGGCCAGGGCGATCTGGAGGCCCTTGGTGTGTGATGGTGCAC 736

QY 749 TTCAGGCTGTGTGTCTGTGGGGTGTGTGGGGCCCTGTGTGACAAGATGCATCCCTGGAG 808

DB 737 TCCAGGGCATCATCTCTGGGGTCT---AGACCCCTGTGGGAGGTCCGACAAACCTGGGG 793

QY 809 TCTACACCTATTTTGAAGTATGTGGACTGGATCCGGATGATCATGAGGAACAATGTAC 868

DB 794 TCTATACCAACATCTGCGGTACCTGTGACTGGATCAAGAGATCATAGGCAGCAAGGGCT 853

QY 869 CT 870

DB 854 CT 855

RESULT 14

US-09-070-526-1

; Sequence 1, Application US/09070526

; Patent No. 6100059

; GENERAL INFORMATION:

; APPLICANT: SOUTHAN, CHRISTOPHER

; APPLICANT: CLINKENBEARD, HELEN

; APPLICANT: BURGESS, NICOLA

; TITLE OF INVENTION: No. 6100059el Compounds

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RATNER & PRESTIA

; STREET: P.O. BOX 980

; CITY: VALLEY FORGE

; STATE: PA

; COUNTRY: USA

; ZIP: 19482

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/070,526

; FILING DATE: 30-APR-1998

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9711952.3

; FILING DATE: 9-JUN-1997

; APPLICATION NUMBER: EP 9730964.4

; FILING DATE: 1-DEC-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: PRESTIA, PAUL F

; REGISTRATION NUMBER: 23,031

; REFERENCE/DOCKET NUMBER: GH-30353

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-407-0700

; TELEFAX: 610-407-0701

; TELEX: 846169

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 944 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-09-070-526-1

Query Match 16.9%; Score 184.8; DB 3; Length 944;

Best Local Similarity 55.6%; Pred. No. 2.3e-41;

Matches 395; Conservative 1; Mismatches 308; Indels 6; Gaps 2;

QY 157 GGGCTCAGCCAGGACCCACACCGAGATTTCAATGCACTGAGTGTGGGGTAACCA 216

DB 174 GGACACTCCAGGGCCACAGAGGACAAAGGTGCTGGGGGTCTATGAGTGCCACCCCATTCG 233

QY 217 CAGCCGTGGCAGGTGGGGCTGTTTTCAGGGCCACCGCTCGCGTGGGGGGTGTCTCTATT 276

DB 234 CAGCCTTGGCAGCGGCTTGTTCAGGGCCAGCAATCTCTGTGGGGTGTCTCTGTA 293

QY 277 GACCAAGGTGGTCTCTCAGCGGCTCATGTCAGCGGCGAGGAGTCTGGGTGGCGCTG 336

DB 294 GGTGGCAACTGGGTCTTACAGTCCCACTGTAAAAAACCGAAATACACAGTACGCGCTG 353

QY 337 GGGGACACAGCTCAGCCAGCTCCACTGGACCGGAGAGATCCGGCAGCGGCTTCTCT 396

DB 354 GGAGACCAAGCTTACAGATAAGATGGCCAGAGCAAGAAATACCTGTGGTTCAGTCC 413

QY 397 GTGACCCATCCGGCT---ACCTGGAGCCTCGAGAGCCAGCAGCAGACCTCCGGGTG 453

DB 414 ATCCACACCCCTGCTATATAACAGAGCGGATGTGGAGGACCAACACCATGATCTGATGCTT 473



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 05:53:50 ; Search time 4625 Seconds  
(without alignments)  
10224.270 Million cell updates/sec

Title: US-10-015-385A-193

Perfect score: 1091

Sequence: 1 caagcaggcatccctctgg.....aggggcaaaaaaaaaa 1091

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.man.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1091	100.0	1091	6	AX697125	Sequence
2	1091	100.0	1091	9	AY358524	Homo sapi
3	498	45.6	811	9	HSMB08731	EX648580 Homo sapi
4	487	44.6	619	6	AR263961	Sequence
5	440	40.3	840	6	AR337568	Sequence
6	440	40.3	840	6	BD130047	Human ser
7	370	33.9	9120	9	AF164623	Homo sapi
8	370	33.9	11820	9	AF135025	Homo sapi
c 9	370	33.9	132323	9	AC011473	Homo sapi
c 10	370	33.9	230000	9	AF243527	Homo sapi
c 11	249	22.8	142334	2	AC073185	Homo sapi
c 12	227	20.8	405	6	AX070938	Sequence
c 13	217	19.9	174724	2	AC140096	Pan trogl
c 14	217	19.9	200792	2	AC130782	Pan trogl
c 15	127	11.6	142334	2	AC073185	Homo sapi
c 16	75	6.9	176647	2	AC130188	Papio anu
c 17	26	2.4	938	4	AB100595	Canis fam
c 18	25	2.3	25	6	AX697374	Sequence
c 19	24	2.2	24	6	AX697375	Sequence
c 20	24	2.2	110000	2	AC119867_2	Continuation (3 of
c 21	24	2.2	154879	9	AL450104	Human DNA
c 22	24	2.2	159391	2	AC027113	Homo sapi
c 23	24	2.2	167302	9	AC099756	Homo sapi
c 24	24	2.2	175572	2	AC025349	Homo sapi
c 25	24	2.2	184895	5	AL844197	Zebrafish
c 26	24	2.2	186971	9	AP003501	Homo sapi
c 27	24	2.2	226890	2	AC135541	Rattus no
c 28	24	2.2	228413	2	AC127853	Rattus no
c 29	24	2.2	243655	2	AC099172	Rattus no
c 30	23	2.1	578	11	G78591	S209P6196PD
c 31	23	2.1	786	6	AX375740	Sequence
c 32	23	2.1	944	6	AX375744	Sequence
c 33	23	2.1	954	4	PIGELAI	Sus scrofa
c 34	23	2.1	959	4	SSELS1	Sus scrofa
c 35	23	2.1	970	6	E00424	Pig elastas
c 36	23	2.1	1020	6	AX823252	Sequence
c 37	23	2.1	1130	6	AR256996	Sequence
c 38	23	2.1	1606	6	AX676264	Sequence
c 39	23	2.1	1613	6	AR256990	Sequence
c 40	23	2.1	1718	9	AF536382	Homo sapi
c 41	23	2.1	1718	9	BC062334	Homo sapi
c 42	23	2.1	1755	9	BC036846	Homo sapi
c 43	23	2.1	110000	2	AC079517_1	Continuation (2 of
c 44	23	2.1	159242	8	AP002902	Oryza sat
c 45	23	2.1	159440	8	AP002900	Oryza sat

# ALIGNMENTS

RESULT 1	AX697125	Sequence 193 from Patent WO0078961.	1091 bp	DNA	linear	PAT 02-APR-2003
AX697125	LOCUS	AX697125				
AX697125	DEFINITION	Sequence 193 from Patent WO0078961.				
AX697125	ACCESSION	AX697125				
AX697125	VERSION	AX697125.1	GI:29498087			
AX697125	KEYWORDS					
AX697125	SOURCE	Homo sapiens (human)				
AX697125	ORGANISM	Homo sapiens				
AX697125	REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.				
AX697125	AUTHORS	Ferrara,N., Stewart,T.A., Williams,P.M., Baker,K.P., Desnoyers,L., Eaton,D.L., Gao,W.Q., Pan,J., Botstein,D., Fong,S., Goddard,A., Godowski,P.J., Gurney,A.L., Smith,V., Tumas,D., Wood,W.I.,				

Grimaldi, C.J., Hillan, K.J., Paoni, N.F., Roy, M.A. and Watanabe, C.K.  
Secreted and transmembrane polypeptides and nucleic acids encoding  
the same  
Patent: WO 0078961-A 193 28-DEC-2000;  
Genentech Inc. (US)  
Location/Qualifiers  
1. .1091  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1091; DB 6; Length 1091;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCAGGTCATCCCTTGGTACCTTCAAGAGAAAGCAGAGGCGCAGAGGTGGGGGC 60  
DB 1 CAAGCAGGTCATCCCTTGGTACCTTCAAGAGAAAGCAGAGGCGCAGAGGTGGGGGC 60  
QY 61 ACAGGAAAGGTCAGCTTGGATTCCTTCCCGCAGACTTTGGAGTGACCCACC 120  
DB 61 ACAGGAAAGGTCAGCTTGGATTCCTTCCCGCAGACTTTGGAGTGACCCACC 120  
QY 121 ATGGGGCTCAGCATCTTTTGGTCTCTGTCTTGGCTCAGCGCAGCAGCAGCAGC 180  
DB 121 ATGGGGCTCAGCATCTTTTGGTCTCTGTCTTGGCTCAGCGCAGCAGCAGCAGC 180  
QY 181 AAGATTTTCAATGGGCTAGTGTGGGCTAATCAACAGCGTGGCAGTGGGGGTGTT 240  
DB 181 AAGATTTTCAATGGGCTAGTGTGGGCTAATCAACAGCGTGGCAGTGGGGGTGTT 240  
QY 241 GAGGGCAGCAGCTCGCTCGGGGTGTCCTTATGACACAGGTGGCTCTCACAGG 300  
DB 241 GAGGGCAGCAGCTCGCTCGGGGTGTCCTTATGACACAGGTGGCTCTCACAGG 300  
QY 301 GGTCACTGACGGCAGCAGGTAAGTGTGGCTGCGCTGGGGGAACACAGCCTCAGCAGCTC 360  
DB 301 GGTCACTGACGGCAGCAGGTAAGTGTGGCTGCGCTGGGGGAACACAGCCTCAGCAGCTC 360  
QY 361 GACTGGACGAGCAGATCGGCACAGCGGCTTCTGTACCCATCCCGGTACCTGGGA 420  
DB 361 GACTGGACGAGCAGATCGGCACAGCGGCTTCTGTACCCATCCCGGTACCTGGGA 420  
QY 421 GCTGTCAGCAGGACACAGCAGCAGCTCCGCTGCTGCGCTGCGCTGCGCTGCGCTG 480  
DB 421 GCTGTCAGCAGGACACAGCAGCAGCTCCGCTGCTGCGCTGCGCTGCGCTGCGCTG 480  
QY 481 ACAGCAGGCTTCAACCTTGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 540  
DB 481 ACAGCAGGCTTCAACCTTGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 540  
QY 541 CAGCTCTAGGCTGGGCGATCAACACACACACACACACACACACACACACACACAC 600  
DB 541 CAGCTCTAGGCTGGGCGATCAACACACACACACACACACACACACACACACACAC 600  
QY 601 TGCCTCAACCTTCCATCGTCTCCATGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 660  
DB 601 TGCCTCAACCTTCCATCGTCTCCATGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 660  
QY 661 ACAGCAGCAGTGTGTGTCAGCGCGCTGCGGGGAGAGTGTGTGTGTGTGTGTGTGT 720  
DB 661 ACAGCAGCAGTGTGTGTCAGCGCGCTGCGGGGAGAGTGTGTGTGTGTGTGTGTGT 720  
QY 721 GGGGGCCCCCTGCTGTGGGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780  
DB 721 GGGGGCCCCCTGCTGTGGGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780  
QY 781 CCTGTGGACAGATGGCATCCCTGGAGTCTACACCTATATTGCAAGTATGTGACTGG 840  
DB 781 CCTGTGGACAGATGGCATCCCTGGAGTCTACACCTATATTGCAAGTATGTGACTGG 840  
QY 841 ATCCGGATGATCATGAGGACACTGACCTGTTCTGTCACCTCCACCCCAACCCCTTAA 900

841 ATCCGGATGATCATGAGGACAACTGACCTGTTCTCCACCTCCACCCCAACCCCTTAA 900  
QY 901 CTTGGGTACCCCTCTGGCCCTCAGAGCACCATAATATCTCTCCATCATTCCCTAGCTCC 960  
DB 901 CTTGGGTACCCCTCTGGCCCTCAGAGCACCATAATATCTCTCCATCATTCCCTAGCTCC 960  
QY 961 ACTCTGTGGCTGGGAACTTTCTTGGAACTTTAACTCTCTGCGACCCCTTTTAAAGACCA 1020  
DB 961 ACTCTGTGGCTGGGAACTTTCTTGGAACTTTAACTCTCTGCGACCCCTTTTAAAGACCA 1020  
QY 1021 CGAGCGGGGTGAGAACTGTCATAGTCTGATATTAATTAATTAATTAATTAATTAATTA 1080  
DB 1021 CGAGCGGGGTGAGAACTGTCATAGTCTGATATTAATTAATTAATTAATTAATTAATTA 1080  
QY 1081 AAAAAA 1091  
DB 1081 AAAAAA 1091

RESULT 2  
AY358524  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

AY358524  
Homo sapiens clone DNA65409 KLK12 (UNQ669) mRNA linear PRI 03-OCT-2003  
AY358524  
AY358524.1 GI:37182170  
FLJ\_CDNA.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1091)  
Clark, R.F., Gurney, A.L., Abaya, E., Baker, K., Baldwin, D., Brush, J.,  
Chen, J., Chow, B., Chui, C., Crowley, C., Currell, B., Deuel, B.,  
Dowd, P., Eaton, D., Foster, J., Grimaldi, C., Gu, Q., Hass, P.E.,  
Heldens, S., Huang, A., Kim, H.S., Klimowski, L., Jin, Y., Johnson, S.,  
Lee, J., Lewis, L., Liao, D., Mark, M., Robbie, E., Sanchez, C.,  
Schoenfeld, J., Seshagiri, S., Simmons, L., Singh, J., Smith, V.,  
Stinson, J., Vagstad, A., Vandien, K., Watanabe, C., Weiland, D., Woods, K.,  
Xie, M.H., Yansura, D., Yi, S., Yu, G., Yuan, J., Zhang, M., Zhang, Z.,  
Zhang, Z., Zou, W.I. and Godowski, P.  
The Secreted Protein Discovery Initiative (SPDI), a Large-Scale  
Effort to Identify Novel Human Secreted and Transmembrane Proteins:  
A Bioinformatics Assessment  
Genome Res. 13 (10), 2265-2270 (2003)  
12975309  
2 (bases 1 to 1091)  
Clark, R.F.  
Direct Submission  
Submitted (01-AUG-2003) Department of Bioinformatics, Genentech,  
Inc., 1 DNA Way, South San Francisco, CA 94080, USA  
location/Qualifiers  
1. .1091  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DNA65409"  
1. .1091  
/locus\_tag="UNQ669"  
121. .867  
/locus\_tag="UNQ669"  
/note="PRO1303"  
/codon\_start=1  
/product="KLK12"  
/protein\_id="AAQ88888.1"  
/db\_xref="GI:37182171"  
/translation="MGLSIFLLLCVLGSAATPKIFNGTECGRNSQFQWVGLPEGTS  
LRCGVLDHRWLTAHCSGRYWRIGHSLSOLDWTEQIRHSQSVTHPGVGLGAS  
TSHEVLDRLRLRLPVRVTSVQPLPLPNCATAGTCHVSGWGTINHPNPFDDLIQ  
CLNLSVSHATCHGVYFGRITSNMVGAGVPGQDACQSDSGPLVCGGLVQGLVSWGS  
VGPCQDGIPIGVYTIICKYVDWIRNMRNN"

FEATURES  
source  
PUBMED  
AUTHORS  
JOURNAL  
TITLE  
JOURNAL  
FEATURES  
source  
gene  
CDS  
ORIGIN

[illegible][illegible]



```
QY 878 CCACCTCCACCCGCCACCCCTTAACTTGGGTACCCCTCTGGGCCCTCAGAGCACCATAATCT 937
Db 526 CCACCTCCACCCGCCACCCCTTAACTTGGGTACCCCTCTGGGCCCTCAGAGCACCATAATCT 585

QY 938 CTTCCATCAGCTTCCCTAGCTCCACTCTTGTGGCTGGGAACCTTCTTGGAACTTTAACT 997
Db 586 CTTCCATCAGCTTCCCTAGCTCCACTCTTGTGGCTGGGAACCTTCTTGGAACTTTAACT 645

QY 998 CTGCGAGCCCTTCTAAGACCCACAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATA 1057
Db 646 CTGCGAGCCCTTCTAAGACCCACAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATA 705

QY 1058 AATATAAATGAAGGAGGG 1075
Db 706 AATATAAATGAAGGAGGG 723

RESULT 4
AR263961
LOCUS AR263961 619 bp DNA linear PAT 29-JAN-2003
DEFINITION Sequence 139 from patent US 6331427.
ACCESSION AR263961
VERSION AR263961.1 GI:28075965
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 619)
TITLE Robison,K.E.
JOURNAL Protease homologs
FEATURES Patent: US 6331427-A 139 18-DEC-2001;
Location/Qualifiers
source 1..619
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 44.6%; Score 487; DB 6; Length 619;
Best Local Similarity 100.0%; Pred. No. 5.2e-278;
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AGACTTTGGAAGTGACCCACCACTGGGGCTCAGCATCTTTTCTCTCTGTGTCTTCTGGG 159
Db 69 AGACTTTGGAAGTGACCCACCACTGGGGCTCAGCATCTTTTCTCTCTGTGTCTTCTGGG 128

QY 160 CTCAGCCAGGAGCCACACCCGAAGATTTTCAATGGCACTAGTGTGGGGTAACTCAAG 219
Db 129 CTCAGCCAGGAGCCACACCCGAAGATTTTCAATGGCACTAGTGTGGGGTAACTCAAG 188

QY 220 CGTGGCAGGTGGGGCTGTTTCAGGCGCACCCAGCTCGCTGGGGGGTGTCTTATTGAC 279
Db 189 CGTGGCAGGTGGGGCTGTTTCAGGCGCACCCAGCTCGCTGGGGGGTGTCTTATTGAC 248

QY 280 CACAGTGGGTCTCAGACGGCTCACTGCAGCGCAGCAGTACTGGGTGGCGCTGGGG 339
Db 249 CACAGTGGGTCTCAGACGGCTCACTGCAGCGCAGCAGTACTGGGTGGCGCTGGGG 308

QY 340 GAACACAGCTCAGCCAGCTCAGCTGGACCGGAGCAGATCCGCGCAGCGGCTTCTGTG 399
Db 309 GAACACAGCTCAGCCAGCTCAGCTGGACCGGAGCAGATCCGCGCAGCGGCTTCTGTG 368

QY 400 ACCATCCCGGCTACTCGGAGCGCTCGACGAGCAGCAGCTCGGCTGTCTGGG 459
Db 369 ACCATCCCGGCTACTCGGAGCGCTCGACGAGCAGCAGCTCGGCTGTCTGGG 428

QY 460 CTGGCGCTCCCGCTCCGCGTAACAGCAGCTTCAACCCCTGCCCTGCCCAATGACTGT 519
Db 429 CTGGCGCTCCCGCTCCGCGTAACAGCAGCTTCAACCCCTGCCCTGCCCAATGACTGT 488

QY 520 GCAACCGCTGGCAGCGAGTGGCAGCTCTCAGCTGGGGTATCACCACCCACCGGAC 579
Db 489 GCAACCGCTGGCAGCGAGTGGCAGCTCTCAGCTGGGGTATCACCACCCACCGGAC 548
```

```
QY 580 CCATTCC 586
Db 549 CCATTCC 555

RESULT 5
AR337568
LOCUS AR337568 840 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 5 from patent US 6566498.
ACCESSION AR337568
VERSION AR337568.1 GI:33723963
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 840)
TITLE Ni,J. and Ruben,S.M.
JOURNAL Human serine protease and serpin polypeptides
FEATURES Patent: US 6566498-A 5 20-MAY-2003;
Location/Qualifiers
source 1..840
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 40.3%; Score 440; DB 6; Length 840;
Best Local Similarity 99.8%; Pred. No. 5.7e-250;
Matches 490; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 585 CCCGATCTGCTCCAGTGCCCTCAACCTCTCCATCGTCTCCCATGCGCACCTGCCATGGTGT 644
Db 321 CCCGATCTGCTCCAGTGCCCTCAACCTCTCCATCGTCTCCCATGCGCACCTGCCATGGTGT 380

QY 645 GTATCCCGGGAGATCAAGCAACATGTTGTGCGAGGGCGCTCCGGGGCAGGATGC 704
Db 381 GTATCCCGGGAGATCAAGCAACATGTTGTGCGAGGGCGCTCCGGGGCAGGATGC 440

QY 705 CTGCCAGGTGATTTCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 764
Db 441 CTGCCAGGTGATTTCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 500

QY 765 CTGGGGGTCTCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 824
Db 501 CTGGGGGTCTCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 560

QY 825 CAAGTATGTGACTGATGATCCGGATGATCATGAGGAACTGATGCTGTTTCTCCACCTC 884
Db 561 CAAGTATGTGACTGATGATCCGGATGATCATGAGGAACTGATGCTGTTTCTCCACCTC 620

QY 885 CACCCCGACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGAGCACCATAATCTCTCTCCAT 944
Db 621 CACCCCGACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGAGCACCATAATCTCTCTCCAT 680

QY 945 CACTTCCCTAGCTCCACTCTTGTGGCTGGGAACCTTCTTGGAACTTTAACTCTCGCA 1004
Db 681 CACTTCCCTAGCTCCACTCTTGTGGCTGGGAACCTTCTTGGAACTTTAACTCTCGCA 740

QY 1005 GCCCTTCTAAGACCCACAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAATAA 1064
Db 741 GCCCTTCTAAGACCCACAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAATAA 800

QY 1065 ATGAAGGAGGG 1075
Db 801 ATGAAGGAGGG 811

RESULT 6
BD130047
LOCUS BD130047 840 bp DNA linear PAT 18-SEP-2002
DEFINITION Human serine protease and serpin polypeptide.
ACCESSION BD130047
VERSION BD130047.1 GI:23224992
```

```
KEYWORDS      JP 2002502600-A/3.
SOURCE        Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 840).
AUTHORS      Ruben,S.M. and Ni,J.
TITLE        Human serine protease and serpin polypeptide
JOURNAL      Patent: JP 2002502600-A 3 29-JAN-2002;
              HUMAN GENOME SCIENCES INC
COMMENT      OS Homo sapiens (human)
              PN JP 2002502600-A/3
              PD 29-JAN-2002
              PF 04-FEB-1999 JP 2000530597
              PR 06-FEB-1998 US 60/073961
              PI STEVEN M RUBEN, JIAN NI
              PC
              C12N9/64, A61K38/48, A61P7/02, A61P7/04, A61P7/06, A61P29/00, A61P31/
              00,
              PC A61P35/00, A61P37/06, C07K14/81, C07K16/40, C12N1/15, C12N1/19, PC
              C12N1/21,
              PC C12N5/10, C12N15/09, C12Q1/37, A61K37/547, C12N5/00, C12N15/00 CC
              Human serine protease and serpin polypeptide
              PH Key Location/Qualifiers
              FT CDS (115)..(603).
              Location/Qualifiers
              1..840
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"

FEATURES             source
    source
    Join 40.3%; Score 440; DB 6; Length 840;
    Best Local Similarity 99.8%; Pred. No. 5.7e-250;
    Matches 490; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 595 CCGGATCTGCTCCAGTGCCTCAACTCTCCATCGTCTCCATGCGACCTGCGCATGGTGT 644
DB |||||
DB 321 CCGGATCTGCTCCAGTGCCTCAACTCTCCATCGTCTCCATGCGACCTGCGCATGGTGT 380
QY 645 GTATCCCGGGAGAAATACAGAGCAACATGCTGTGTGTCAGCGCGCGCTCCGGGCGCAGATGC 704
DB |||||
DB 381 GTATCCCGGGAGAAATACAGAGCAACATGCTGTGTGTCAGCGCGCGCTCCGGGCGCAGATGC 440
QY 705 CTGCCAGGGTGATCTGCGGGGCCCTGCTGTGTGTGGGGAGTCTTCAAGGTCTGCTGTC 764
DB |||||
DB 441 CTGCCAGGGTGATCTGCGGGGCCCTGCTGTGTGTGGGGAGTCTTCAAGGTCTGCTGTC 500
QY 765 CTGGGGGTCTGCGGGGCCCTGTCGACAGATGCGCATCCCTGGAGTCTACACCTATATTG 824
DB |||||
DB 501 CTGGGGGTCTGCGGGGCCCTGTCGACAGATGCGCATCCCTGGAGTCTACACCTATATTG 560
QY 825 CAAATGATGTGAATCGGATCCGGATGATCATAGGAACAACACTGCTGTTCTCCACCTC 884
DB |||||
DB 561 CAAATGATGTGAATCGGATCCGGATGATCATAGGAACAACACTGCTGTTCTCCACCTC 620
QY 885 CACCCCGACCCCTTAATTGGGTACCCCTCTGGCCCTCAGACACCAATATCTCTCCAT 944
DB |||||
DB 621 CACCCCGACCCCTTAATTGGGTACCCCTCTGGCCCTCAGACACCAATATCTCTCCAT 680
QY 945 CACTTCCCTAGCTCCACTCTGTTGGCTGGGACTCTTCTGGAATTTAACTCTGCA 1004
DB |||||
DB 681 CACTTCCCTAGCTCCACTCTGTTGGCTGGGACTCTTCTGGAATTTAACTCTGCA 740
QY 1005 GCGCTTCTAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAAATAA 1064
DB |||||
DB 741 GCGCTTCTAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAAATAA 800
QY 1065 ATGAAGGAGGG 1075
DB |||||
DB 801 ATGAAGGAGGG 811
```

```
AF164623      9120 bp      DNA      linear      PRI 26-JUN-2000
Homo sapiens trypsin-like serine protease (TLSP) gene, complete
cds.
ACCESSION    AF164623
VERSION      AF164623.1
KEYWORDS     GI:5713130
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 9120)
AUTHORS      Yousef,G.M., Scorilas,A. and Diamandis,E.P.
TITLE        Genomic organization, mapping, tissue expression, and hormonal
              regulation of trypsin-like serine protease (TLSP PRSS20), a new
              member of the human kallikrein gene family
              Genomics 63 (1), 88-96 (2000)
JOURNAL      20130117
MEDLINE      10662548
PUBMED
REFERENCE     2 (bases 1 to 9120)
AUTHORS      Yousef,G.M., Scorilas,A. and Diamandis,E.P.
TITLE        Direct Submission
JOURNAL      Submitted (01-JUL-1999) Pathology and Laboratory Medicine, Mount
              Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1x5,
              Canada
FEATURES             Location/Qualifiers
    Location/Qualifiers
    1..9120
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="19"
    /map="19q13.3-q13.4"
    2313..7622
    /gene="TLSP"
    join(2313..2398,4189..4263,5061..5217,5545..5810,
    5627..6763,7158..7622)
    /gene="TLSP"
    /product="trypsin-like serine protease"
    join(4224..4263,5061..5217,5545..5810,6627..6763,
    7158..7310)
    /gene="TLSP"
    /note="Kallikrein-like serine protease"
    /codon_start=1
    /product="trypsin-like serine protease"
    /protein_id="AAD47815.1"
    /db_xref="GI:5713131"
    /translation="MRILQLILLALATGLVGGETRIIKGFCKPHSQPWQALPEKTR
    LLCCATLAPRWLLTAHCLKPRYIVHLQHNLOKEGEQTRTATESFPHPFNNSL
    PNKDRNDIMLVKMASPVSIWAVRPLTSSRCVATGTSCLISGWSSTSPOLRPHPT
    LRCANITIEHOKCENAYPGNITDTWCASVQEGGKDCQGDGSGGLVCLNQLQIIS
    WQDPCATITRPGYVTKYKVDVMTQENKN"

Query Match      33.9%; Score 370; DB 9; Length 9120;
Best Local Similarity 100.0%; Pred. No. 3.4e-208;
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 708 CCAGGTGATCTGGGGGCCCTGCTGTGTGTGGGGAGTCTTCAAGTCTGGTCTCTG 767
DB |||||
DB 394 CCAGGTGATCTGGGGGCCCTGCTGTGTGTGGGGAGTCTTCAAGTCTGGTCTCTG 453
QY 768 GGGGTCTCTGGGGGCCCTGCTGGACAAGATGGCATCCCTGGAGTCTACACCTATATTGCAA 827
DB |||||
DB 454 GGGGTCTCTGGGGGCCCTGCTGGACAAGATGGCATCCCTGGAGTCTACACCTATATTGCAA 513
QY 828 GTATGTGACTGGATCCGATGATCATGAGAACTGACCTGTTTCTCCACCTCCAC 887
DB |||||
DB 514 GTATGTGACTGGATCCGATGATCATGAGAACTGACCTGTTTCTCCACCTCCAC 573
QY 888 CCCACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGAGCACCATATCTCTCCATCAC 947
DB |||||
DB 574 CCCACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGAGCACCATATCTCTCCATCAC 633
```

Qy	948	TTCCCTAGTCTCACCTCTTTGGTGGCTGGGAACCTTCTTGGAACTTTAACTCCTGCCAGCC	1007
Db	634	TTCCCTAGTCTCACCTCTTTGGTGGCTGGGAACCTTCTTGGAACTTTAACTCCTGCCAGCC	693
Qy	1008	CTTCTAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTCGAATAAATAAATG	1067
Db	694	CTTCTAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTCGAATAAATAAATG	753
Qy	1068	AAGAGGGGC	1077
Db	754	AAGAGGGGC	763
RESULT 8			
LOCUS	AF135025	11820 bp	linear
DEFINITION	Homo sapiens kallikrein-like protein 5 gene, alternative splice products, complete cds.		
ACCESSION	AF135025		
VERSION	AF135025.2	GI:6166248	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE 1	Antkarcner Res. 79, 2843-2852 (1999)		
REFERENCE 2	Diemandis, E.P., Yousef, G.M., Luo, L.Y., Magklara, A. and Obiezu, C.V. The new human kallikrein gene family: implications in carcinogenesis Trends Endocrinol. Metab. 11 (2), 54-60 (2000)		
REFERENCE 3	Yousef, G.M., Magklara, A. and Diamandis, E.P. Kik-L5 is a novel serine protease and a new member of the human kallikrein gene family-differential expression in breast cancer Unpublished		
REFERENCE 4	Yousef, G.M., Magklara, A., Scorilas, A. and Diamandis, E.P. Cloning of new alternatively spliced forms of the kallikrein-like gene 5 (KLK-L5) Unpublished		
REFERENCE 5	Yousef, G.M., Magklara, A. and Diamandis, E.P. Direct Submission Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada		
REFERENCE 6	Yousef, G.M., Magklara, A. and Diamandis, E.P. Direct Submission Submitted (01-NOV-1999) Pathology and Laboratory Medicine, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada		
REFERENCE 7	Sequence update by submitter On Nov 1, 1999 this sequence version replaced gi:4589276. Location/Qualifiers		
COMMENT	1. .11820		
FEATURES	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/chromosomes="19"		
	/map="19q13.3-q13.4"		
	join(3743..3840,3995..4050,4496..4655,6500..6759,7714..7847,9178..9543)		
	/product="kallikrein-like protein 5"		
	join(3743..3840,3995..4050,4496..4655,6500..6759,7714..7847,9178..9543)		
mRNA			
mRNA			

Db	9534	AAGGAGGGC	9543	
RESULT 9	AC011473	132323 bp	DNA	linear
LOCUS	AC011473/c			
DEFINITION	Homo sapiens chromosome 19, BAC BC349142 (CTC-518B2), complete sequence.			
ACCESSION	AC011473			
VERSION	AC011473			
KEYWORDS	HTG			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	1 (bases 1 to 132323)			
AUTHORS	Lamerdin, J.E., McCready, P.M., Skowronski, E., Viswanathan, V., Burkhardt-Schultz, K., Gordon, L., Dias, J., Ramirez, M., Stilwagen, S., Phan, H., Velasco, N., Do, L., Regala, W., Terry, A., Brower, A., Barnes, J., Danganan, L., Erler, A., Christensen, M., Georgescu, A., Avila, J., Liu, S., Andreise, T., Trankheim, M., Attix, C., Amico-Keller, G., Coefield, J., Duarte, S., Lucas, S., Bruce, R., Thomas, P., Quan, G., Krommiller, B., Arellano, A., Sanders, C., Ow, D., Nolan, M., Trong, S., Kobayashi, A., Olsen, A.S. and Carrano, A.V.			
TITLE	Sequence analysis of chromosome 19q13.4			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 132323)			
AUTHORS	DOE Joint Genome Institute.			
TITLE	Direct Submission			
JOURNAL	Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA			
REFERENCE	3 (bases 1 to 132323)			
AUTHORS	Lamerdin, J.E.			
TITLE	Direct Submission			
JOURNAL	Submitted (12-OCT-2000) DOE Joint Genome Institute, Lawrence Livermore National Laboratory, Livermore, CA 94550			
COMMENT	On Oct 12, 2000 this sequence version replaced gi:7458725. Map and sequence oriented from centromere to q-telomere. BC349142 (CTC-518B2) overlaps BAC BC85745 (CTB-147C22, AC011483) on the left from bases 1 to 2,226 of this accession, and overlaps BAC BC892989 (CTD-3187P9, AC063977) on the right from bases 120,922 to 132,323 of this accession. Additional chromosome 19 map and sequence information are available at: <a href="http://www-bio.lnl.gov/bbrp/genome/genome.html">http://www-bio.lnl.gov/bbrp/genome/genome.html</a> . Location/Qualifiers			
FEATURES	1..132323			
source	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9606"			
	/chromosome="19"			
	/map="19q13.4"			
	/clone="CTC-518B2 (LIML clone name BC349142)"			
	/complement(<77..1783)			
	/gene="KLK8"			
gene	/note="kallikrein 8 (neuropsin/ovasin); synonyms: TADG14, PRSS19, HNP; protease, serine, 19 (neuropsin/ovasin)"			
	/complement(join(<77..339,505..664,1179..1256,1613..1783))			
	/gene="KLK8"			
	/product="KLK8 mRNA"			
	/complement(join(<77..339,505..664,1179..1248))			
	/gene="KLK8"			
	/note="potential molecular marker for ovarian carcinomas"			
	/codon_start=1			
	/product="kallikrein 8 (neuropsin/ovasin) [amino acids 1-164]"			
	/protein_id="AAG23254.1"			
	/db_xref="GI:10799393"			
	/translation="MGRPRRAAKTWMFLLLGGAWAGHSRAQEDKVLGGHEGCPHSQ PWQALFQGLCGVIGVGNVLTAAHCKPKYVRLGDHSLQNKDGPQEIPIVMQ SIHPHCYNSDVEDHNDLMLQLRDQASLGSVKVPISLADHCTPGQKCTVSGGTV TSPR"			
	382..473			
repeat_region				

repeat_region	/rpt_family="MSR1"			
	/complement(1374..1554)			
	/rpt_family="MSR1"			
repeat_region	1723..1754			
	/rpt_family="TCC)n"			
repeat_region	/complement(1791..2238)			
	/rpt_family="MSR1"			
gene	/complement(2594..9715)			
	/gene="KLK9"			
	/note="kallikrein 9; synonyms: KLK-L3, kallikrein-like protein 3"			
mRNA	/complement(join(2594..3341,3785..3921,6539..6804,9264..9420,9586..9715))			
	/gene="KLK9"			
	/product="KLK9 mRNA"			
	/complement(join(3192..3341,3785..3921,6539..6804,9264..9420,9586..9628))			
CDS	/gene="KLK9"			
	/note="kallikrein-like protein 3 KLK-L3"			
	/codon_start=1			
	/product="kallikrein 9"			
	/protein_id="AAG23255.1"			
	/db_xref="GI:10799394"			
	/translation="MKGLICALLSLLAGHWADTRAGBRCRNSOPWQAGLPHLT RLFCGATLISDRWLLTAACHCKPKYLWRLGHEHLWKWGPQLFRVTFFPHPGNKD LSANDNDIMLRIPROARLSPAVQPLNLSOTCVSPGQCCLISGWAVSSPRLFPV TLOCANISILENKLCHWAYPGHISDSMLCAGLWEGGSGSCQDGGPLVCNGTLAGV SGGAEPSCRPRPAVYTSVCHYLDWIQIMEN"			
repeat_region	3406..3496			
	/rpt_family="MSR1"			
repeat_region	3558..3700			
	/rpt_family="MSR1"			
repeat_region	4042..4139			
	/rpt_family="GA-rich"			
repeat_region	4140..4449			
	/rpt_family="AluSx"			
repeat_region	4450..4483			
	/rpt_family="GA-rich"			
repeat_region	4529..4812			
	/rpt_family="AluSc"			
repeat_region	/complement(5043..5253)			
	/rpt_family="Charlie8"			
repeat_region	5370..5480			
	/rpt_family="AluJo"			
repeat_region	5605..5712			
	/rpt_family="L3"			
repeat_region	5878..6014			
	/rpt_family="AluSg"			
repeat_region	6015..6043			
	/rpt_family="(TAA)n"			
repeat_region	6044..6235			
	/rpt_family="AluSg"			
repeat_region	6302..6340			
	/rpt_family="CT-rich"			
repeat_region	6346..6411			
	/rpt_family="MER104"			
repeat_region	6422..6520			
	/rpt_family="GA-rich"			
repeat_region	7221..7263			
	/rpt_family="(TG)n"			
repeat_region	7840..7938			
	/rpt_family="L3"			
repeat_region	8130..8448			
	/rpt_family="AluSg"			
repeat_region	8634..8919			
	/rpt_family="MER33"			
repeat_region	10629..10758			
	/rpt_family="L2"			
repeat_region	10755..10905			
	/rpt_family="L2"			
repeat_region	10976..11096			
	/rpt_family="MIR"			
repeat_region	11097..11128			

```
repeat_region      /rpt_family=" (TTTG)n"
repeat_region      complement(11133..11411)
repeat_region      /rpt_family="Alusg"
repeat_region      11412..11515
repeat_region      /rpt_family="MIR"
repeat_region      11680..11979
repeat_region      /rpt_family="Alusx"
repeat_region      complement(11989..12069)
repeat_region      /rpt_family="MIR"
repeat_region      complement(12070..12390)
repeat_region      /rpt_family="AluJo"
repeat_region      complement(12391..12498)
repeat_region      /rpt_family="MIR"
repeat_region      12503..12606
repeat_region      /rpt_family="L2"
repeat_region      complement(12694..12976)
repeat_region      /rpt_family="L2"
repeat_region      12993..13101
repeat_region      /rpt_family="L2"
repeat_region      13287..13393
repeat_region      /rpt_family="L2"
repeat_region      complement(13507..13804)
repeat_region      /rpt_family="Alusq"
repeat_region      complement(14351..14379)
repeat_region      /gene="KlK10"
repeat_region      /note="Kallikrein 10; synonyms: NES1, PRSSL1"
repeat_region      complement(join(14351..15033,15498..15631,15963..16237,
repeat_region      17191..17371,19124..19220,19708..19779))
repeat_region      /gene="KlK10"
repeat_region      /product="KlK10 mRNA"
repeat_region      14384..14530
repeat_region      /standard_name="stSG30247"
repeat_region      /note="RH48452, WIAF-3662-STS"
repeat_region      complement(14432..14604)
repeat_region      /rpt_family="MIR"
repeat_region      complement(join(14881..15033,15498..15631,15963..16237,
repeat_region      17191..17371,19124..19220))
repeat_region      /gene="KlK10"
repeat_region      /note="protease, serine-like, 1"
repeat_region      /codon_start=1
repeat_region      /product="kallikrein 10"
repeat_region      /protein_id="AAG23256.1"
repeat_region      /db_xref="GI:10799395"
repeat_region      /translation="MRPHLHLSAARALAKLPLIMAOAAEALLPQNDRLD
repeat_region      PEATGSPCARQSPQVPSLFLGSHFCAGVLVDQSVLTAHCKNKLWARVGDHLL
repeat_region      LLOEQELRRTRVVPKHYHOGSGPILPRRDEHDLMLKLARPVVLGPRVRLQLEPV
repeat_region      RCAOPGQCOVAGMGTAAARVKNKGLTCSITILSPKECEVYFPGVWNNICAGL
repeat_region      DRGQDPCQSDSGGFLVCDTELQGLISWGVYFCGSAQHPAVYTIQKTMWINKVIRSN
repeat_region      "
repeat_region      15245..15420
repeat_region      /rpt_family=" (TTGGGG)n"
repeat_region      15715..15967
repeat_region      /rpt_family="MSR1"
repeat_region      complement(16473..16818)
repeat_region      /rpt_family="Alusq"
repeat_region      17787..17912
repeat_region      /rpt_family="AluJo"
repeat_region      17913..17932
repeat_region      /rpt_family=" (CAA)n"

Query Match      33.9%; Score 370; DB 9; Length 132323;
Best Local Similarity 100.0%; Pred. No. 3.5e-208;
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      708      CCAGGGTGAATCTGGGGGCCCCCTGGTGTGGGGAGTCTCTTCAAGGTCTGGTCTCTG 767
Db      29542      CCAGGGTGAATCTGGGGGCCCCCTGGTGTGGGGAGTCTCTTCAAGGTCTGGTCTCTG 29483
QY      768      GGGGTCTGGGGCCCTGTGACAAAGATGGATCCCTGGAGTCTACACTATATTGCAA 827
Db      29482      GGGGTCTGGGGCCCTGTGACAAAGATGGATCCCTGGAGTCTACACTATATTGCAA 29423
QY      828      GTATGTGGACTGGATCCGGATCATGAGGAACTAGTCTTTCTCCCTCCACTCCAC 887
```

```
Db      29422      GTATGTGGACTGGATCCGGATCATGAGGAACTAGTCTTTCTCCCTCCACTCCAC 29363
QY      888      CCCACACCCCTTAACCTTGGGTACCCCTCTGGCCCTCAGAGCACCACCAATATCTCTCCATCAC 947
Db      29362      CCCACACCCCTTAACCTTGGGTACCCCTCTGGCCCTCAGAGCACCACCAATATCTCTCCATCAC 29303
QY      948      TTCCTAGCTCCACTCTTGTGTGGCTGGGAACTTTTGGAACTTTAACTCTCTCCAGCC 1007
Db      29302      TTCCTAGCTCCACTCTTGTGTGGCTGGGAACTTTTGGAACTTTAACTCTCTCCAGCC 29243
QY      1008      CTTCTAAGACCCACGAGCGGGTGAGAGAACTGTGCAATAGTCTGGATAATAATAATG 1067
Db      29242      CTTCTAAGACCCACGAGCGGGTGAGAGAACTGTGCAATAGTCTGGATAATAATAATG 29183
QY      1068      AAGGAGGGGC 1077
Db      29182      AAGGAGGGGC 29173

RESULT 10
AF243527/c
LOCUS      AF243527      230000 bp      DNA      linear      PRI 21-NOV-2000
DEFINITION      Homo sapiens serine protease gene cluster, complete sequence.
ACCESSION      AF243527
VERSION      AF243527.1      GI:11244757
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 230000)
AUTHORS      Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,
Moss, P., Paepfer, B. and Wang, K.
TITLE      Sequencing and expression analysis of the serine protease gene
cluster located in chromosome 19q13 region
JOURNAL      Gene 257 (1), 119-130 (2000)
MEDLINE      20510030
PubMed      11054574
REFERENCE      2 (bases 1 to 230000)
AUTHORS      Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,
Moss, P., Paepfer, B. and Wang, K.
TITLE      Direct Submission
JOURNAL      Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE,
Bothell, WA 98021, USA
FEATURES
source
1..230000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13"
/complement(<6963..>11517)
/gene="KLK1"
complement(join(<6963..7118,7668..7804,7923..8212,
9482..9641,11472..>11517))
/gene="KLK1"
/product="renal kallikrein"
complement(join(6963..7118,7668..7804,7923..8212,
9482..9641,11472..11517))
/gene="KLK1"
/note="serine protease"
/codon_start=1
/product="renal kallikrein"
/protein_id="AAG3353.1"
/db_xref="GI:11244758"
/translation="MMFLVLCIALSLGGTGAAPPIQSRIVGVWECHQSOPQWALYH
FSTFGCGILVHRQWLTAHCISDNVQLWLRHNLFDNDENTAQFVHVSESPHGFN
MSLENTHTROADEYDSDMLMLLTPADTITDAVKVELPTEPVSGTCLASGGS
IEPENSEPDDLQCVDLKLLPNDCEKKAHVQKVTDFMLCVGHLEGGDKTCVGS
MCDGVLGQVTSWGIYPCGTPKPSVAVRVLVSIVYKWIETIAENS"
complement(join(<13552..13704,14377..14530,14678..14917,
15416..15569,19204..>19246))
mRNA
```

```

CDS
/complement (join(13552..13704,14377..14530,14678..14917,
15416..15569,19204..19246))
/notes="serine protease"
/codon_start=1
/product="ACO protease"
/protein_id="AAG33354.1"
/db_xref="GI:11244759"
/translation="MWLLLTSLFLASTAAQDQKLEGEBCAPHSPQWVALYERGR
FNCASLISPHWLVLSAAHCSRMVRVRLGHNLRKDGPRQLRTSVIHPHYEARS
HRNDIMLLRLVQARLNPQVRPAVLPRCHPGEACVSGGLVPLSPFVSLSDTLHC
ANTSISDTSCKSPYRLNTMVCAGRGRAESCEGSGGLVCGGILQGIQVSGWD
VPCDNTKPGVYTKVCHYLEWIRETKRN"
<42595..>.47769
/gene="KLK3"
mRNA
join(<42595..42640,43880..44039,45669..45955,46099..46235,
47614..47769)
/gene="KLK3"
CDS
/product="prostate specific antigen"
/join(42595..42640,43880..44039,45669..45955,46099..46235,
47614..47769)
/notes="KLK3"
/codon_start=1
/product="prostate specific antigen"
/protein_id="AAG33355.1"
/db_xref="GI:11244760"
/translations="MVPVVFLLTSVTWIGAAPLILSRVGVGWECEKHSOPWQVIVAS
RGRACGVVLHPQWVLTAAHCLRNKSVLLGRHSILFHPEDTGVQVSHSFHPPLYD
MSLLKHQSLRPDSDSHDMLLRLSPAEITDAKVMDLPTQSPALGTTCYASGWSGI
EPREFITPKLQCVLDHIVSNDCAQHPQKVTKFMCLAGRTGKGSKTCSGDSGGPLV
CNGVLQGITSWGSEPCALPERPSLYTKVHYRWKIDTIANP"
<61139..>.66229
/gene="KLK2"
mRNA
join(<61139..61184,62391..62550,64142..64428,64542..64678,
66074..66229)
/gene="KLK2"
CDS
/product="glandular kallikrein"
/join(61139..61184,62391..62550,64142..64428,64542..64678,
66074..66229)
/notes="KLK2"
/codon_start=1
/product="glandular kallikrein"
/protein_id="AAG33356.1"
/db_xref="GI:11244761"
/translations="MMDLVLSIALSVGCTGAVPLIOSRIQVGVGWECEKHSOPWQVAVS
HGAHCGVVLHPQWVLTAAHCLRNKSVQVWLGRHNLFEPEDTQORYPVSHSPHPLYN
MSLLKHQSLRPDSDSHDMLLRLSPAKITDVYKVLGLPTQBPALGTTCYASGWSGI
EPEEFILRPSLQCVSHLSLNDMCARAYSEKVTFFMLCAGLWTGGKDTCCGDSGGPLV
CNGVLQGITSWGSEPCALPEKPAVYTKVHYRWKIDTIANP"
complement (join(<94602..94754,96027..96163,96247..96497,
96919..97081,98345..>.98405))
/product="serine protease"
complement (join(94602..94754,96027..96163,96247..96497,
96919..97081,98345..>.98405))
/codon_start=1
/product="serine protease"
/protein_id="AAG33357.1"
/db_xref="GI:11244762"
/translations="WATACNPWGLYLLVAGSLVSGSCQIINGDCSPQWPW
QALVWENELFCGVVLHPQWVLSAAHCFQNSYITLGLHLSLEADQEPSSQWVEASLS
VRHPYENRPLANDMLIKLSDVSSESTIRISASQCTFAGNSCLVSGWGLANGR
VMTLQCVNVSVEEVSKLYDPLYPHSMFCAGGHHQDKSCNGDSGGPLICNGYLQ
GLVSGFAPCGQGVGVVTLCKFTFETKTVQAS"
complement (<131301..>.140303)
/gene="KLK5"
/notes="synonym: SCCE"
complement (join(<131301..131456,136310..136443,
136529..136785,137525..137690,140255..>.140303))
/gene="KLK5"
/product="stratum corneum trypsin-like serine protease"
complement (join(131301..131456,136310..136443,

```

```

136529..136785,137525..137690,140255..140303))
/gene="KLK5"
/codon_start=1
/product="stratum corneum trypsin-like serine protease"
/protein_id="AAG33358.1"
/db_xref="GI:11244763"
/translations="MVLICALITALLGVTDARSDSSSRILNGSDCDMTQWQAAL
LLRPNQLYCGAVLVHFWQLMTAAHCRKRVFRLGHYSLSPVYESGQGFQGVKSIPH
PGVSHGHNDMLIKINRIRPTKVRPNVNSCHPDSAGTKCLVSGWGTTKGFVHE
KYLQCLNLSVLSCKECEDYPRQIDDTMFCAGDKAGRDCQDGSQGVVVCNGSLQGL
VSMGDYPCARPNRPGVYTNLCKFTKTIQTIQANS"
complement (join(<146834..146986,149628..149764,
151186..151433,155052..155208,155948..>.155987))
/product="protease M"
complement (join(146834..146986,149628..149764,
151186..151433,155052..155208,155948..155987))
/notes="serine protease; also called neurosin or zyme"
/codon_start=1
/product="protease M"
/protein_id="AAG33359.1"
/db_xref="GI:11244764"
/translations="MKKLMVLSLIAAAWAEQKLVHGGPCDKTSHPYQAALYTSQH
LLCGVLIHPLWLTAAHCKKNLQFLGKHNLRQESSQSSVAVVHPDYDAAS
HDQIMLLRLAPKLSLQPLERDCSANTTSCHILGWKGTADGDRPDTIQCAVI
HLVSRCECHAYPGQITONMLCAGDEKYGKSCQDGSQGVVLCGDHLRGLVSGNIFC
GSKEKPGVYTNVCTYTNWIKTIQAK"
complement (<165420..>.170283)
/gene="KLK7"
/notes="synonym: SCCE"
complement (join(<165420..165575,167672..167808,
168124..168371,169651..169798,170211..>.170283))
/gene="KLK7"
/product="stratum corneum chymotryptic enzyme"
complement (join(165420..165575,167672..167808,
168124..168371,169651..169798,170211..>.170283))
/notes="KLK7"
/codon_start=1
/product="stratum corneum chymotryptic enzyme"
/protein_id="AAG33360.1"
/db_xref="GI:11244765"
/translations="MARSLILPLIQLLSLALETAGEEAQDKIIDGAPCARGSHPWQ
VALLSGNLHCGVVLNWRWLTAAHCKKNVTVHLGSDTLGDRRAORTKAKSRPHQ
GYSTOYVNDMLKLNLSQARLSMVYKVLPSRCEPCTCTVSCWGTTSPTDVFPP
SLDMCVDVKLJSPQDCTKYVDLLENSMLCAGIPSKKNACNGSDSGPLVCRGTLQGL
VSWGTFPCQPNPDGVTYVCAFTKWINITMKHR"
complement (join(<183943..184098,185635..185768,
187865..188127,188293..188452,188967..>.189036))
/product="neuropsin"
complement (join(183943..184098,185635..185768,
187865..188127,188293..188452,188967..>.189036))
/notes="serine protease; also called ovasin"
/codon_start=1
/product="neuropsin"
/protein_id="AAG33361.1"
/db_xref="GI:11244766"
/translations="MGRPRPRAKTMWMLILGGAWAGHRAQERKVLGGHECQPHSQ
PMQAALFQGOQLLCGVLVGNWVLTAAHCKKPKYTVRLGDSLNQKGPQEIIPVQ
STPHCYNSDVEDHNDMLLQKCRDAYSGLSKVKPISLADHCTQPGQKCTVSGWTV
TSPRFNPTDLNCAEVKIPQKQEDAPYQGITDMGVACAGSKSGADTCQDGSQGVPLV
DGAQGITSWGSDPCGRSDKPGVYTNICRYLDWIKIISGK"
complement (join(<190980..191129,191573..191709,
194324..194589,197048..197204,197370..>.197412))
/product="kallikrein-like 3"
complement (join(190980..191129,191573..191709,
194324..194589,197048..197204,197370..>.197412))
/notes="serine protease"
/codon_start=1
/product="kallikrein-like 3"
/protein_id="AAG33362.1"
/db_xref="GI:11244767"
/translations="MKLLGALLSLLAGHWADTRATCABECRPNQSQWQAGLPHLT
RLFCGATLISDRWLTAAHCRKRVLWVRLGHHLWKWEGPQLPRVTFFHPGPNKD

```

LSANDHDDIMLIRLPQARLSPAVOPINLSQTCVSPGMQCLISGWAIVSPKALFPV  
TLQANISILNKUICHWAYPHIHSIDMLCAGLWEGRGSCQDGGPLVNGTLAGVV  
SGGAPCSRPFRPAVYTSVCHYLDIWOIEMEN"

Query Match 33.9%; Score 370; DB 9; Length 230000;

Best Local Similarity 100.0%; Pred. No. 3.6e-208; Mismatches 0; Indels 0; Gaps 0;

Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 708 CCAGGCTGATCTGGGGGCCCTGGTGTGTGGGGAGTCTTCAAGTCTGGTCTG 767

Db 217325 CCAGGCTGATCTGGGGGCCCTGGTGTGTGGGGAGTCTTCAAGTCTGGTCTG 217266

QY 768 GGGGTCTGTGGGGCCCTGTGGACAGATGGCATCCTGGAGTCTACACTATATTGCA 827

Db 217265 GGGGTCTGTGGGGCCCTGTGGACAGATGGCATCCTGGAGTCTACACTATATTGCA 217206

QY 828 GTATGTGACTGGATCCGGATGATCATGAGGAACAACCTGCTTCTCCACCTCCAC 887

Db 217205 GTATGTGACTGGATCCGGATGATCATGAGGAACAACCTGCTTCTCCACCTCCAC 217146

QY 888 CCCGACCCCTTAACCTGGTACCCCTCTGGCCCTCAGACACCAATATCTCTCCATCAC 947

Db 217145 CCCGACCCCTTAACCTGGTACCCCTCTGGCCCTCAGACACCAATATCTCTCCATCAC 217086

QY 948 TTCCCTAGCTCCACTCTTTGGCTGGGAACCTTTTGGAACTTTAACTCTCTGCCGCC 1007

Db 217085 TTCCCTAGCTCCACTCTTTGGCTGGGAACCTTTTGGAACTTTAACTCTCTGCCGCC 217026

QY 1008 CTTCTAAGACCCACGAGCGGGGTGAGAGTGTGCAATAGTCTGGAATAATAATAATG 1067

Db 217025 CTTCTAAGACCCACGAGCGGGGTGAGAGTGTGCAATAGTCTGGAATAATAATAATG 216966

QY 1068 AAGGAGGGGC 1077

Db 216965 AAGGAGGGGC 216956

#### RESULT 11

AC073185/c 142334 bp DNA linear HTG 17-AUG-2000  
LOCUS AC073185 Homo sapiens chromosome 19 clone RP11-10111, WORKING DRAFT  
DEFINITION AC073185  
ACCESSION AC073185  
VERSION AC073185.4 GI:9838034  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Waterston,R.H.  
The sequence of Homo sapiens clone

1 (bases 1 to 142334)

Waterston,R.H.  
Unpublished  
2 (bases 1 to 142334)

Waterston,R.H.  
Direct Submission

Submitted (10-JUN-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On Aug 17, 2000 this sequence version replaced gi:9653152.

#### COMMENT

Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc/index.shtml>

Center project name: H.NH001011

Center project name: H.NH001011

Center project name: H.NH001011

Center project name: H.NH001011

Center project name: H.NH001011

Center project name: H.NH001011

Center project name: H.NH001011

Center project name: H.NH001011

Center project name: H.NH001011

Center project name: H.NH001011

Center project name: H.NH001011

Center project name: H.NH001011

Center project name: H.NH001011

Consensus quality: bases at least Q40  
Consensus quality: bases at least Q30  
Consensus quality: bases at least Q20  
Insert size: 135000; agarose-fp  
Insert size: 135000; sum-of-contigs  
Quality coverage: 3.86 in Q20 bases; agarose-fp  
Quality coverage: 3.87 in Q20 bases; sum-of-contigs  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 32 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1452: contig of 1452 bp in length

1453: gap of unknown length

1553: contig of 1185 bp in length

2737: gap of unknown length

2837: contig of 1828 bp in length

4665: gap of unknown length

4765: contig of 1289 bp in length

6054: gap of unknown length

6154: contig of 1637 bp in length

6155: gap of unknown length

7791: contig of 2135 bp in length

7891: gap of unknown length

10028: contig of 1769 bp in length

10127: gap of unknown length

11895: contig of 3075 bp in length

11896: gap of unknown length

15070: contig of 3075 bp in length

15071: gap of unknown length

15171: contig of 3306 bp in length

18477: gap of unknown length

18577: contig of 2961 bp in length

21537: gap of unknown length

21637: contig of 2988 bp in length

24626: gap of unknown length

24726: contig of 2858 bp in length

27583: gap of unknown length

27684: contig of 3209 bp in length

30893: gap of unknown length

30993: contig of 4217 bp in length

35209: gap of unknown length

35310: contig of 3952 bp in length

39261: gap of unknown length

39361: contig of 2995 bp in length

42356: gap of unknown length

42457: contig of 3965 bp in length

46421: gap of unknown length

46521: contig of 4636 bp in length

46522: gap of unknown length

51158: contig of 6159 bp in length

51258: gap of unknown length

57417: contig of 4194 bp in length

57517: gap of unknown length

61711: contig of 4439 bp in length

61811: gap of unknown length

66249: contig of 5631 bp in length

66250: gap of unknown length

66350: contig of 5374 bp in length

71980: gap of unknown length

72081: contig of 6764 bp in length

77455: gap of unknown length

77554: contig of 4231 bp in length

84319: gap of unknown length

84419: contig of 5670 bp in length

88650: gap of unknown length

88749: contig of 5646 bp in length

94420: gap of unknown length

94519: contig of 7069 bp in length

100166: gap of unknown length

100266: contig of 7659 bp in length

107335: gap of unknown length

107434: contig of 7659 bp in length

115093: contig of 7659 bp in length





Db 91 TTTTCCCCCAGACTTGGAGAGTGAACCCACACCGAGGGGCTCAGCATCTTTTGGCTCCTGTGT 150  
 Qy 151 GTTCTTGGGCTAGCCAGGAGGACACACCGAAGATTTTCAATGSCACTGAGTGTGGGCGT 210  
 Db 151 GTTCTTGGGCTAGCCAGGAGGACACACCGAAGATTTTCAATGSCACTGAGTGTGGGCGT 210  
 Qy 211 AACTCAGCGGTGGCAGGTGGGCTGTTTGAAGGACACCGCTGCGCTGCGGGGGTGTTC 270  
 Db 211 AACTCAGCGGTGGCAGGTGGGCTGTTTGAAGGACACCGCTGCGCTGCGGGGGTGTTC 270  
 Qy 271 CTATTACACACAGGTGGGTCTCTCACAGCGGCTCACTGCAGCGGCAG 317  
 Db 271 CTATTACACACAGGTGGGTCTCTCACAGCGGCTCACTGCAGCGGCAG 317

RESULT 13  
 AC140096/c  
 LOCUS 174724 bp DNA linear HTG 14-JUN-2003  
 DEFINITION Pan troglodytes clone RP43-63K10, WORKING DRAFT SEQUENCE, 10  
 ordered pieces.  
 AC140096  
 AC140096.4 GI:31745211  
 VERSION HTG; HTGS PHASE2; HTGS DRAFT.  
 KEYWORDS Pan troglodytes (chimpanzee)  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS  
 Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariaga, X.,  
 Chu, G., Coleman, B., Coleman, H., Engle, J., Granite, S., Guan, X.,  
 Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Hu, P.,  
 Hurlb, B., Iqbal, R., Karlins, S., Kwong, P., Latic, P., Lee-fan, S.-Q.,  
 Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masiello, C.,  
 Maskeri, B., McDowell, J., Pagnirigan, C., Pearson, R., Portnoy, M.E.,  
 Prasad, A., Reddix-Bugue, N., Schandler, K., Schueler, M.G., Shah, K.,  
 Sison, C., Stantrypop, S., Thomas, J.W., Thomas, P.J., Tsipouri, V.,  
 Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.  
 NISC Comparative Sequencing Initiative  
 Unpublished  
 2 (bases 1 to 174724)  
 Green, E.D.  
 Direct Submission  
 Submitted (21-FEB-2003) NTH Intramural Sequencing Center, 8717  
 Grovenmont Circle, Gaithersburg, MD 20877, USA  
 3 (bases 1 to 174724)  
 Green, E.D.  
 Direct Submission  
 Submitted (14-JUN-2003) NIH Intramural Sequencing Center, 8717  
 Grovenmont Circle, Gaithersburg, MD 20877, USA  
 On Jun 14, 2003 this sequence version replaced gi:29725783.

COMMENT  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: <http://www.nisc.nih.gov>  
 Contact: nisc.zoonhgr@nih.gov  
 ----- Project Information  
 Center project name: dxz  
 Center clone name: 063K10

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 172841 bases at least Q40  
 Consensus quality: 173490 bases at least Q30  
 Consensus quality: 173737 bases at least Q20  
 Insert size: 150000; agarose-Ep  
 Insert size: 173824; sum-of-contigs  
 Quality coverage: 15.91x in Q20 bases; agarose-fp  
 Quality coverage: 13.73x in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 10 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.

\* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

1 47271 47270: contig of 47270 bp in length  
 \* 47371 58115: contig of 10745 bp in length  
 \* 58116 61050: contig of 2835 bp in length  
 \* 61051 65346: contig of 4196 bp in length  
 \* 65347 71536: contig of 6090 bp in length  
 \* 71537 83139: contig of 11503 bp in length  
 \* 83140 85133: contig of 1894 bp in length  
 \* 85134 88937: contig of 3604 bp in length  
 \* 88938 129108: contig of 40171 bp in length  
 \* 129109 129208: contig of 45516 bp in length.  
 \* 129209 174724: contig of 45516 bp in length.

## FEATURES

## source

1. 174724  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /clone\_lib="RP43-63K10"  
 /clone\_lib="RP43"

## misc\_feature

1. 79098  
 /note="clone overlaps with GenBank Accession Number  
 AC130782 clone CH251-355A20 (center project name dhz)"  
 1. 47270  
 /note="assembly\_fragment  
 clone\_end:17"

## misc\_feature

1. 47270  
 /note="assembly\_fragment  
 vector\_side:left"

## misc\_feature

47371..58115  
 /note="assembly\_fragment"

## misc\_feature

58216..61050  
 /note="assembly\_fragment"

## misc\_feature

61151..65346  
 /note="assembly\_fragment"

## misc\_feature

65447..71536  
 /note="assembly\_fragment"

## misc\_feature

71637..83139  
 /note="assembly\_fragment"

## misc\_feature

83240..85133  
 /note="assembly\_fragment"

## misc\_feature

85234..88837  
 /note="assembly\_fragment"

## misc\_feature

88938..129108  
 /note="assembly\_fragment"

## misc\_feature

113257..174724  
 /note="clone overlaps with GenBank Accession Number  
 AC133392 clone CH251-126O24 (center project name dxa)"  
 129209..174724

On Nov 22, 2002 this sequence version replaced gi:22218452.

COMMENT

/note="assembly\_fragment  
clone\_end:SP6  
vector\_side:right"

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: http://www.nisc.nih.gov

Contact: nisc.zoo@hgrl.nih.gov

Project Information

Center project name: dhz

Center clone name: 355A20

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies. The low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 196656 bases at least Q40  
Consensus quality: 197883 bases at least Q30  
Consensus quality: 198879 bases at least Q20  
Insert size: 215000; agarose-fp  
Quality coverage: 9.03x in Q20 bases; agarose-fp  
Quality coverage: 9.73x in Q20 bases; sum-of contigs

\* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

\* 1 28306: contig of 28306 bp in length  
\* 28307 28406: gap of unknown length  
\* 28407 37856: contig of 9450 bp in length  
\* 37857 37957: gap of unknown length  
\* 37957 73522: contig of 35566 bp in length  
\* 73522 83567: contig of 9945 bp in length  
\* 83567 83668: gap of unknown length  
\* 83668 88817: contig of 5150 bp in length  
\* 88817 125611: contig of 36694 bp in length  
\* 125611 159879: contig of 34168 bp in length  
\* 159879 159979: gap of unknown length  
\* 159979 174698: contig of 14719 bp in length  
\* 174698 174799: gap of unknown length  
\* 174799 186382: contig of 11584 bp in length  
\* 186382 186482: gap of unknown length  
\* 186482 193344: contig of 6862 bp in length  
\* 193344 193445: gap of unknown length  
\* 193445 199363: contig of 5919 bp in length  
\* 199363 199464: gap of unknown length  
\* 199464 200792: contig of 1329 bp in length.

Location/Qualifiers  
1. .200792  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="CH251-355A20"  
/clone\_lib="CH251"

FEATURES  
source

ORIGIN  
Query Match 19.9%; Score 217; DB 2; Length 174724;  
Best local similarity 99.2%; Pred. No. 6.5e-117;  
Matches 367; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 708 CCAGGTTGATCTGGGGCCCCCTGGTGTGTGGGGAGTCCCTCAAGGTCTGCTCTG 767  
DB 38225 CCAGGTTGATCTGGGGCCCCCTGGTGTGTGGGGAGTCCCTCAAGGTCTGCTCTG 38166  
QY 768 GGGGTCTGGGGCCCCCTGGGACAGATGGCATCCCTGGAGTCTACACTATATTGCA 827  
DB 38165 GGGGTCTGGGGCCCCCTGGGACAGATGGCATCCCTGGAGTCTACACTATATTGCA 38106  
QY 828 GATGTGGACTGGATCCGGATCATGATGAGGAACTGACCTGTTTCTCCACCTCCAC 887  
DB 38105 GATGTGGACTGGATCCGGATCATGATGAGGAACTGACCTGTTTCTCCACCTCCAC 38046  
QY 898 CCCACCCCTTAACCTGGTACCCCTGGCCCTCAGACACAAATATCTCTCCATCAC 947  
DB 38045 CCCACCCCTTAACCTGGTACCCCTGGCCCTCAGACACAAATATCTCTCCATCAC 37986  
QY 948 TTCCTAGCTCCACTCTTGTGGCTGGAACTTCTTGGAACTTTAACTCTGCCAGCC 1007  
DB 37985 TTCCTAGCTCCACTCTTGTGGCTGGAACTTCTTGGAACTTTAACTCTGCCAGCC 37926  
QY 1008 CTTCTAAGACCCAGCGGGGTGAGAGAGTGTGCAATAGTCTGGAATAATAAATG 1067  
DB 37925 CTTCTAAGACCCAGCGGGGTGAGAGAGTGTGCAATAGTCTGGAATAATAAATG 37866  
QY 1068 AAGGAGGGGC 1077  
DB 37865 AAGGAGGGGC 37856

RESULT 14  
AC130782/c  
LOCUS  
DEFINITION  
Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12 ordered pieces.  
AC130782  
AC130782.2 GI:25167101  
VERSION  
HIG; HTGS\_PHASE2; HTGS\_DRAFT.  
KEYWORDS  
SOURCE  
Pan troglodytes (chimpanzee)  
ORGANISM  
Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
REFERENCE  
1 (bases 1 to 200792)  
Akhtar, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Carriaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laxic, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B., Marquis, E.H., Masello, C., Maskeri, B., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C., Stadtripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 200792)  
Green, E.D.  
Direct Submission  
Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717 Groveomont Circle, Gaithersburg, MD 20877, USA  
3 (bases 1 to 200792)  
Green, E.D.  
Direct Submission  
Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717 Groveomont Circle, Gaithersburg, MD 20877, USA

misc\_feature 1. 28306  
/note="assembly\_fragment  
vector\_end:SP6  
vector\_side:left"  
misc\_feature 28407. 37856  
/note="assembly\_fragment"  
misc\_feature 37957. 73522  
/note="assembly\_fragment"  
misc\_feature 73623. 83567  
/note="assembly\_fragment"  
misc\_feature 93668. 88817  
/note="assembly\_fragment"  
misc\_feature 88918. 125611  
/note="assembly\_fragment"  
misc\_feature 125712. 159879  
/note="assembly\_fragment"  
misc\_feature 159980. 174698  
/note="assembly\_fragment"  
misc\_feature 174799. 186382  
/note="assembly\_fragment"  
misc\_feature 186483. 193344  
/note="assembly\_fragment"  
misc\_feature 193445. 199363  
/note="assembly\_fragment"  
misc\_feature 199464. 200792  
/note="assembly\_fragment  
clone\_end:r7  
vector\_side:right"

## ORIGIN

Query Match 19.9%; Score 217; DB 2; Length 200792;  
Best Local Similarity 99.2%; Pred. No. 6.5e-117;  
Matches 367; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 708 CCAGGATGATTCGGGGCCCCCTGGTGTGGGGAGTCCTTCAAGGTCTGGTCTCTG 767  
Db 148657 CCAGGATGATTCGGGGCCCCCTGGTGTGGGGAGTCCTTCAAGGTCTGGTCTCTG 148598  
QY 768 GGGGTCTGTGGGGCCCTGTGGACAGATGGCATCCTGGAGCTACACCTATTTGCAA 827  
Db 148597 GGGGTCTGTGGGGCCCTGTGGACAGATGGCATCCTGGAGCTACACCTATTTGCAA 148538  
QY 828 GPATGTGACTGGATCCGGATGATCATGAGAACACTGACCTGTTCCTCCACCTCCAC 887  
Db 148537 GPATGTGACTGGATCCGGATGATCATGAGAACACTGACCTGTTCCTCCACCTCCAC 148478  
QY 888 CCCACCCCTTAACTTGGTGTACCCCTCTGGCCCTCAGACACCAATATCTCTCCATCAC 947  
Db 148477 CCCACCCCTTAACTTGGTGTACCCCTCTGGCCCTCAGACACCAATATCTCTCCATCAC 148418  
QY 948 TTCCTCTAGCTCCACTCTTGTGGCTGGAACTTCTTGGAACTTTAACTCTCTGCGAGCC 1007  
Db 148417 TTCCTCTAGCTCCACTCTTGTGGCTGGAACTTCTTGGAACTTTAACTCTCTGCGAGCC 148358  
QY 1008 CTTCTAAGACCCACGAGCGGGGTGAGAGAGTGTGCAATAGTCTGCAATAAATAAATG 1067  
Db 148357 CTTCTAAGACCCACGAGCGGGGTGAGAGAGTGTGCAATAGTCTGCAATAAATAAATG 148298  
QY 1068 AAGGAGGGGC 1077  
Db 148297 AAGGAGGGGC 148288

## RESULT 15

AC073185  
LOCUS AC073185 142334 bp DNA linear HTG 17-AUG-2000  
DEFINITION Homo sapiens chromosome 19 clone RP11-10111, WORKING DRAFT  
SEQUENCE, 32 unordered pieces.  
AC073185  
ACCESSION AC073185  
VERSION AC073185.4 GI:9838034  
KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 142334)  
Waterston,R.H.  
The sequence of Homo sapiens clone  
Unpublished  
2 (bases 1 to 142334)  
Waterston,R.H.  
Direct Submission  
Submitted (10-JUN-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On Aug 17, 2000 this sequence version replaced gi:9653152.

## COMMENT

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site:http://genome.wustl.edu/gsc/index.shtml  
----- Project Information -----  
Center project name: H.NH0010111  
----- Summary Statistics -----  
Sequencing vector: M13; 100%  
Chemistry: Dye-primer ET; 100% of reads  
Chemistry: Dye-terminator Big Dye; 0% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: bases at least Q40  
Consensus quality: bases at least Q30  
Consensus quality: bases at least Q20  
Insert size: 135000; agarose-fp  
Insert size: 139234; sum-of-contigs  
Quality coverage: 3.86 in Q20 bases; agarose-fp  
Quality coverage: 3.87 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 32 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1452: contig of 1452 bp in length  
1453 1552: gap of unknown length  
1553 2737: contig of 1185 bp in length  
2738 2837: gap of unknown length  
2838 4665: contig of 1828 bp in length  
4666 4765: gap of unknown length  
4766 6054: contig of 1289 bp in length  
6055 6154: gap of unknown length  
6155 7791: contig of 1637 bp in length  
7792 7891: gap of unknown length  
7892 10026: contig of 2135 bp in length  
10027 10126: gap of unknown length  
10127 11895: contig of 1769 bp in length  
11896 11995: gap of unknown length  
11996 15070: contig of 3075 bp in length  
15071 15170: gap of unknown length  
15171 18476: contig of 3306 bp in length  
18477 18576: gap of unknown length  
18577 21537: contig of 2961 bp in length  
21538 21637: gap of unknown length  
21638 24625: contig of 2388 bp in length  
24626 24725: gap of unknown length  
24726 27583: contig of 2858 bp in length  
27584 27683: gap of unknown length  
27684 30892: contig of 3209 bp in length  
30893 30992: gap of unknown length  
30993 35209: contig of 4217 bp in length  
35210 35309: gap of unknown length  
35310 39261: contig of 3952 bp in length  
39262 39361: gap of unknown length  
39362 42356: contig of 2995 bp in length



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 04:42:21 ; Search time 509 Seconds  
(without alignments)  
9105.666 Million cell updates/sec

Title: US-10-015-385A-193  
Perfect score: 1091  
Sequence: 1 caagcaggtcatcccttg.....aggggcaaaaaaaaaa 1091

Scoring table: OLIGO NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size: 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001as.\*  
5: Geneseq2001bs.\*  
6: Geneseq2002s.\*  
7: Geneseq2003as.\*  
8: Geneseq2003bs.\*  
9: Geneseq2003cs.\*  
10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1091	100.0	1091	3	Aaa77671 Human PRO
2	1091	100.0	1091	3	Aac58114 Human PRO
3	1091	100.0	1091	3	Aaa37075 Human PRO
4	1091	100.0	1091	4	Aaf54341 DNA encod
5	1091	100.0	1091	8	Acd68378 Novel hum
6	1091	100.0	1091	8	Ach04480 Human CDN
7	1091	100.0	1091	8	AcD68024 Novel hum
8	1091	100.0	1091	9	AdC18062 Human PRO
9	1091	100.0	1091	9	Add70708 Human CDN
10	1091	100.0	1091	9	Add39785 Human CDN
11	1091	100.0	1091	9	Add70231 Human CDN
12	1091	100.0	1091	9	Add38352 Human CDN
13	1091	100.0	1091	9	Add39308 Human CDN
14	1091	100.0	1091	9	Add38831 Human CDN
15	1091	100.0	1091	9	Add40262 Human CDN
16	1091	100.0	1091	9	Ades0483 Human CDN
17	1091	100.0	1091	9	Ade20095 Human CDN
18	1091	100.0	1091	9	Ade5006 Human CDN
19	1091	100.0	1091	9	Ade21564 Human CDN
20	963	88.3	1185	4	Aah98653 Human EST
21	727	66.6	945	6	Abk94900 Human nov
22	696	63.8	747	7	Aal59914 Human kal
23	487	44.6	619	6	Abk30369 Human G-p

24	440	40.3	840	2	AAx80906	Rax80906 Human CDN
25	370	33.9	11820	3	AAa95944	Aaa95944 Human KIK
26	227	20.8	405	5	AAf65654	Aaf65654 Novel hum
27	60	5.5	60	6	ABn45198	Abn45198 Human spl
28	50	4.6	50	6	ABz04830	Abz04830 Human leu
29	25	2.3	25	3	AAC58071	Aac58071 Human PRO
30	25	2.3	25	4	Aaf54518	Aaf54518 Probe #63
31	25	2.3	25	8	ACd68557	AcD68557 Novel hum
32	25	2.3	25	8	ACH04659	ACH04659 Human sec
33	25	2.3	25	8	ACD68203	ACd68203 Novel hum
34	25	2.3	25	9	ADCI8311	AdCI8311 Human PRO
35	25	2.3	25	9	ADD70957	AdD70957 Human PRO
36	25	2.3	25	9	ADD40034	AdD40034 Human PRO
37	25	2.3	25	9	ADD70480	AdD70480 Human PRO
38	25	2.3	25	9	ADD38601	AdD38601 Human PRO
39	25	2.3	25	9	ADD39557	AdD39557 Human PRO
40	25	2.3	25	9	ADD39080	AdD39080 Human PRO
41	25	2.3	25	9	ADD40511	AdD40511 Human PRO
42	25	2.3	25	9	ADE50732	AdE50732 Human PRO
43	25	2.3	25	9	ADE20344	AdE20344 Human PRO
44	25	2.3	25	9	ADE50255	AdE50255 Human PRO
45	25	2.3	25	9	ADE21813	AdE21813 Human PRO

ALIGNMENTS

RESULT 1  
AAA77671

ID AAA77671 standard; cDNA; 1091 BP.

AC AAA77671;

DT 07-NOV-2000 (first entry)

DE Human PRO1303 cDNA sequence SEQ ID NO:202.

KW Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;  
diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;  
KW angiogenic; proliferative; cardiac; cardiovascular; antiatherosclerotic;  
KW cytosolic; gene therapy; vaccine; ss.

OS Homo sapiens.

PN WO200032221-A2.

XX 08-JUN-2000.

PF 30-NOV-1999; 99WO-US028313.

PR 01-DEC-1998; 98WO-US025108.

PR 16-DEC-1998; 98US-0112850P.

PR 12-JAN-1999; 99US-0115554P.

PR 08-MAR-1999; 99WO-US005028.

PR 12-MAR-1999; 99US-0123957P.

PR 28-APR-1999; 99US-0131445P.

PR 14-MAY-1999; 99US-0134287P.

PR 02-JUN-1999; 99WO-US012252.

PR 23-JUN-1999; 99US-0141037P.

PR 20-JUL-1999; 99US-0144758P.

PR 26-JUL-1999; 99US-0145698P.

PR 01-SEP-1999; 99WO-US020111.

PR 08-SEP-1999; 99WO-US020594.

PR 13-SEP-1999; 99WO-US020944.

PR 15-SEP-1999; 99WO-US021090.

PR 15-SEP-1999; 99WO-US021547.

PR 05-OCT-1999; 99WO-US023089.

PR 29-OCT-1999; 99US-0162506P.

XX (GETH ) GENENTECH INC.

PA Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ;

PI Goddard A, Godowski PJ, Gurney AL, Klein RD, Kuo SS, Paoni NF;









```
QY 241 GAGGACACAGCCTGCGCTGCGGGGTGTCCTTATTGACACACAGGTGGGTCTCTCACAGC 300
Db 241 GAGGGACACAGCCTGCGCTGCGGGGTGTCCTTATTGACACACAGGTGGGTCTCTCACAGC 300
QY 301 GCTCACTGACGGCAGCAGGTACTGGGTGCGCTGGGGGAAACACAGCCTCAGCCAGCTC 360
Db 301 GCTCACTGACGGCAGCAGGTACTGGGTGCGCTGGGGGAAACACAGCCTCAGCCAGCTC 360
QY 361 GACTGACCGAGCAGATCCGGACACAGCGCTCTCTGTGACCCATCCCGGTACTCTGGGA 420
Db 361 GACTGACCGAGCAGATCCGGACACAGCGCTCTCTGTGACCCATCCCGGTACTCTGGGA 420
QY 421 GCCTGACGAGCAGCAGCAGCAGCTCGCGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 480
Db 421 GCCTGACGAGCAGCAGCAGCAGCTCGCGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 480
QY 481 ACCAGCAGCTTCAACCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCT 540
Db 481 ACCAGCAGCTTCAACCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCT 540
QY 541 CACGTCTCAGCTGGGCGATCACCAACACACACACACACACACACACACACACACACACAC 600
Db 541 CACGTCTCAGCTGGGCGATCACCAACACACACACACACACACACACACACACACACACAC 600
QY 601 TGCTCAACCTCTCCATCGTCTCCATGCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 TGCTCAACCTCTCCATCGTCTCCATGCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 ACAGGACACATGTTGTGAGGCGGCTGCCGGGCGAGGATGCTGCCAGGGTGAATCT 720
Db 661 ACAGGACACATGTTGTGAGGCGGCTGCCGGGCGAGGATGCTGCCAGGGTGAATCT 720
QY 721 GGGGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 721 GGGGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 781 CCTGTGGACAAGATGCGATCCCTGGAGTCTACACCTATATTGCAAGTATGTGAGTCTG 840
Db 781 CCTGTGGACAAGATGCGATCCCTGGAGTCTACACCTATATTGCAAGTATGTGAGTCTG 840
QY 841 ATCGGATGATCATGAGAACATGACCTGTTTCTCCACCTCCACCCCCACCCCTTAA 900
Db 841 ATCGGATGATCATGAGAACATGACCTGTTTCTCCACCTCCACCCCCACCCCTTAA 900
QY 901 CTTGGGTACCCCTCTGCGCCCTCAGACACCAATATCTCCTCCATCACTTCCCTAGCTCC 960
Db 901 CTTGGGTACCCCTCTGCGCCCTCAGACACCAATATCTCCTCCATCACTTCCCTAGCTCC 960
QY 961 ACTCTGTTGGCTGGAACTTCTTGGAACCTTTAACTCTCCAGCCCTTCTTAAGACCCA 1020
Db 961 ACTCTGTTGGCTGGAACTTCTTGGAACCTTTAACTCTCCAGCCCTTCTTAAGACCCA 1020
QY 1021 CGAGCGGGGTGAGAGAGTGTCAATAGTCTGGAATAATATAAATGAAGAGGGGAAA 1080
Db 1021 CGAGCGGGGTGAGAGAGTGTCAATAGTCTGGAATAATATAAATGAAGAGGGGAAA 1080
QY 1081 AAAAAAAAAA 1091
Db 1081 AAAAAAAAAA 1091
```

## RESULT 4

AAF54341  
ID AAF54341 standard; DNA; 1091 BP.

XX  
AC AAF54341;

DT 02-APR-2001 (first entry)

XX DNA encoding protein of the invention #54.

DE Secreted; transmembrane; gene therapy; ss.

XX

OS  
XX  
PN  
XX  
XX  
PD  
XX  
PF  
XX  
XX  
PR  
PR  
PR  
PR  
PR  
PR  
PR  
PR  
PR  
PR  
PR  
PR  
PR  
XX

Unidentified.  
WO200078961-A1.  
28-DEC-2000.  
18-FEB-2000; 2000WO-US004342.  
23-JUN-1999; 99US-0141037P.  
20-JUL-1999; 99US-0144758P.  
26-JUL-1999; 99US-0145698P.  
01-SEP-1999; 99WO-US020111.  
29-OCT-1999; 99US-0162506P.  
30-NOV-1999; 99WO-US028313.  
02-DEC-1999; 99WO-US028551.  
16-DEC-1999; 99WO-US030095.  
05-JAN-2000; 2000WO-US000219.  
06-JAN-2000; 2000WO-US000376.

(GETH) GENENTECH INC.

PA  
XX  
PI  
PI  
PI  
PI  
XX  
DR  
XX  
XX  
PT  
PT  
PT  
XX  
XX  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
XX  
SQ

Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;  
Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;  
Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;  
Williams PM, Wood WI;  
WPI; 2001-071395/08.

Secreted and transmembrane proteins and nucleic acids designated PRO,  
useful as hybridization probes, in chromosome and gene mapping and gene  
therapy.

Claim 2; Fig 107; 787pp; English.

The present invention relates to secreted and transmembrane proteins.  
These proteins and the DNA encoding them may be used as hybridization  
probes, in chromosome and gene mapping and in the generation of anti-  
sense RNA and DNA. They may also be used to generate either of anti-  
transgenic animals or knockout animals which are in turn useful for  
development and screening of therapeutically useful reagents. The nucleic  
acids may also be used in gene therapy

Sequence 1091 BP; 221 A; 335 C; 313 G; 222 T; 0 U; 0 Other;

Query Match 100.0%; Score 1091; DB 4; Length 1091;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CAAGCAGGTCAATCCCTTGGTGAACCTTCAAGAGAAGCAGAGAGGCGAGAGTGGGGGC 60
Db 1 CAAGCAGGTCAATCCCTTGGTGAACCTTCAAGAGAAGCAGAGAGGCGAGAGTGGGGGC 60
QY 61 ACAGGGAAGGGTGACCTCTGAGATTCCTTTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 120
Db 61 ACAGGGAAGGGTGACCTCTGAGATTCCTTTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 120
QY 121 ATGGGGCTCAGCATCTTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 121 ATGGGGCTCAGCATCTTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 181 AAGATTTTCAATGGCACTGAGTGTGGGGCTAACTCAAGCCGTGGCAGGTGGGGTGT 240
Db 181 AAGATTTTCAATGGCACTGAGTGTGGGGCTAACTCAAGCCGTGGCAGGTGGGGTGT 240
QY 241 GAGGCAACAGCCTCGCTGCGGGGTGCTCTTATGACCAAGTGGGTCTCTCAGCG 300
Db 241 GAGGCAACAGCCTCGCTGCGGGGTGCTCTTATGACCAAGTGGGTCTCTCAGCG 300
QY 301 GCTCACTGACGGCAGCAGGTACTGGGTGCGCTGGGGGAAACACAGCCTCAGCCAGCTC 360
Db 301 GCTCACTGACGGCAGCAGGTACTGGGTGCGCTGGGGGAAACACAGCCTCAGCCAGCTC 360
QY 361 GACTGACCGAGCAGATCCGGACACAGCGCTTCTCTGTGACCCATCCCGGTACTCTGGGA 420
```



```
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-0103314P.
PR 07-OCT-1998; 98US-0103315P.
PR 07-OCT-1998; 98US-0103328P.
PR 07-OCT-1998; 98US-0103395P.
PR 07-OCT-1998; 98US-0103396P.
PR 07-OCT-1998; 98US-0103401P.
PR 08-OCT-1998; 98US-0103633P.
PR 08-OCT-1998; 98US-0103678P.
PR 08-OCT-1998; 98US-0103679P.
PR 08-OCT-1998; 98US-0103711P.
PR 14-OCT-1998; 98US-0104257P.
PR 20-OCT-1998; 98US-0104987P.
PR 20-OCT-1998; 98US-0105000P.
PR 20-OCT-1998; 98US-0105002P.
PR 21-OCT-1998; 98US-0105104P.
PR 22-OCT-1998; 98US-0105169P.
PR 22-OCT-1998; 98US-0105266P.
PR 26-OCT-1998; 98US-0105693P.
PR 26-OCT-1998; 98US-0105694P.
PR 27-OCT-1998; 98US-0105807P.
PR 27-OCT-1998; 98US-0105881P.
PR 27-OCT-1998; 98US-0105882P.
PR 28-OCT-1998; 98US-0106062P.
PR 28-OCT-1998; 98US-0106023P.
PR 28-OCT-1998; 98US-0106029P.
PR 28-OCT-1998; 98US-0106030P.
PR 28-OCT-1998; 98US-0106032P.
PR 28-OCT-1998; 98US-0106033P.
PR 28-OCT-1998; 98US-0106178P.
PR 29-OCT-1998; 98US-0106248P.
PR 29-OCT-1998; 98US-0106384P.
PR 29-OCT-1998; 98US-0108500P.
PR 30-OCT-1998; 98US-0106464P.
PR 03-NOV-1998; 98US-0106856P.
PR 03-NOV-1998; 98US-0106902P.
PR 03-NOV-1998; 98US-0106905P.
PR 03-NOV-1998; 98US-0106919P.
PR 03-NOV-1998; 98US-0106932P.
PR 10-NOV-1998; 98US-0106934P.
PR 17-NOV-1998; 98US-010775P.
PR 17-NOV-1998; 98US-0108779P.
PR 17-NOV-1998; 98US-0108787P.
PR 17-NOV-1998; 98US-0108801P.
PR 17-NOV-1998; 98US-0108802P.
PR 17-NOV-1998; 98US-0108806P.
PR 17-NOV-1998; 98US-0108807P.
PR 17-NOV-1998; 98US-0108857P.
PR 18-NOV-1998; 98US-0108925P.
PR 18-NOV-1998; 98US-0108848P.
PR 18-NOV-1998; 98US-0108849P.
PR 18-NOV-1998; 98US-0108850P.
PR 18-NOV-1998; 98US-0108851P.
PR 18-NOV-1998; 98US-0108852P.
PR 18-NOV-1998; 98US-0108858P.
PR 18-NOV-1998; 98US-0108904P.
PR 22-DEC-1998; 98US-0113296P.
PR 30-DEC-1998; 98US-0114233P.
PR 05-JAN-1999; 99WO-US000106.
PR 16-APR-1999; 99US-0129674P.
PR 23-JUN-1999; 99US-0141037P.
PR 26-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021194.
PR 29-OCT-1999; 99US-0162506P.
PR 30-NOV-1999; 99WO-US028313.
PR 02-DEC-1999; 99WO-US028551.
PR 16-DEC-1999; 99WO-US030095.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.

PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004342.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006884.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 10-DEC-2000; 2000WO-US032678.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 01-JUN-2001; 2001WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 04-SEP-2001; 2001US-00946374.
XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Botstein D, Desnovers L, Eaton DL, Ferrara N, Fong S;
PI Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-585293/55.
DR P-PSDB; ABO33635.
XX
PT Novel isolated PRO polypeptides e.g. PRO1130, PRO1275, PRO1418, PRO1555,
PT PRO1787 that modulate glucose or free fatty acid uptake by skeletal
PT muscle cells, and are useful for treating diabetes, hyper- or hypo-

Query March 100.0%; Score 1091; DB 8; Length 1091;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAAGCAGGTGATCCCTTGGTGACCTTCAAGAGAACAGAGAGGAGAGTGGGGGC 60
Db 1 CAAGCAGGTGATCCCTTGGTGACCTTCAAGAGAACAGAGAGGAGAGTGGGGGC 60
Qy 61 ACAGGGAAGAGGTGACCTCTGAGATCCCTTTTCCCGGAGCTTGGAGTGAAGTGCACCC 120
Db 61 ACAGGGAAGAGGTGACCTCTGAGATCCCTTTTCCCGGAGCTTGGAGTGAAGTGCACCC 120
Qy 121 ATGGGGCTCAGCATCTTTTGGTCTCTGTGTGTTCTTTGGGCTCAGCCAGGAGGACACCG 180
Db 121 ATGGGGCTCAGCATCTTTTGGTCTCTGTGTGTTCTTTGGGCTCAGCCAGGAGGACACCG 180
Qy 181 AAGATTTTCAATGGCACTGAGTGTGGCGGTAACTCAAGCCGTTGGAGTGGGGCTGT 240
Db 181 AAGATTTTCAATGGCACTGAGTGTGGCGGTAACTCAAGCCGTTGGAGTGGGGCTGT 240
Qy 241 GAGGCAACGAGCTGCGCTGCGGGGGTGTCTTTATTTAGACACAGTGGGTCTCTCAGCG 300
Db 241 GAGGCAACGAGCTGCGCTGCGGGGGTGTCTTTATTTAGACACAGTGGGTCTCTCAGCG 300
Qy 301 GCTCACTGACGGGAGCAGAGTACTGGGTGCGCTTGGGGGAAACACAGCCCTCAGCCAGCTC 360
Db 301 GCTCACTGACGGGAGCAGAGTACTGGGTGCGCTTGGGGGAAACACAGCCCTCAGCCAGCTC 360
Qy 361 GACTGACCGAGCAGATCCGACAGCGCTTCTCTGTGACCCATCCCGGCTACCTGGGA 420
Db 361 GACTGACCGAGCAGATCCGACAGCGCTTCTCTGTGACCCATCCCGGCTACCTGGGA 420
Qy 421 GCCTGACGAGCAGCAGACGACCTCCGGCTGTGCGGCTGCGGCTGCGGCTGCGGCTGCGGTA 480
Db 421 GCCTGACGAGCAGCAGACGACCTCCGGCTGTGCGGCTGCGGCTGCGGCTGCGGCTGCGGTA 480
Qy 481 ACCAGAGGTTCAACCCCTGCCCTGCCCAATGACTGTGCAACCGTGGCAGCGAGTGC 540
```





```
Db      601  |||||TGCCCTCAACCTCTCCATGCTCTCCATGCCATGCCATGGTGTATCCCGGGAATC 560
Qy      661  ACAGCAACATGGTGTGTCAGCGCGGTCCTCCCGGCGAGATGCCCTGCCAGGATGATTC 720
Db      661  ACAGCAACATGGTGTGTCAGCGCGGTCCTCCCGGCGAGATGCCCTGCCAGGATGATTC 720
Qy      721  GGGGGCCCCCTGCTGTGTGGGGAGTCCCTTCAAGGTCTGGTGTCTCTGGGGTCTGTGGGG 780
Db      721  GGGGGCCCCCTGCTGTGTGGGGAGTCCCTTCAAGGTCTGGTGTCTCTGGGGTCTGTGGGG 780
Qy      781  CCCTGTGGACAGATGGCATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGCACTGG 840
Db      781  CCCTGTGGACAGATGGCATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGCACTGG 840
Qy      841  ATCCGGATGATCATGAGGAACAACTGACCTGTTCTCCACCTCCACCTCCACCTCCCTTAA 900
Db      841  ATCCGGATGATCATGAGGAACAACTGACCTGTTCTCCACCTCCACCTCCACCTCCCTTAA 900
Qy      901  CTTGGGTACCCCTCTGGCCCTCAGAGCAACATATCTCTCCATCACTTCCCTAGCTCC 960
Db      901  CTTGGGTACCCCTCTGGCCCTCAGAGCAACATATCTCTCCATCACTTCCCTAGCTCC 960
Qy      961  ACTCTGTGTGGCTGGGAACTTCTTGGAACTTAACTCTGCGAGCCCTTCTTAAGACCCA 1020
Db      961  ACTCTGTGTGGCTGGGAACTTCTTGGAACTTAACTCTGCGAGCCCTTCTTAAGACCCA 1020
Qy      1021  CGAGCGGGGTGAGAGAAGTGTCAATAGTCTGGAATAAATAAATGAAGGAGGGGCAAA 1080
Db      1021  CGAGCGGGGTGAGAGAAGTGTCAATAGTCTGGAATAAATAAATGAAGGAGGGGCAAA 1080
Qy      1081  AAAAAAAAAA 1091
Db      1081  AAAAAAAAAA 1091

RESULT 7
ACD68024
ID   ACD68024 standard; cDNA; 1091 BP.
XX
AC   ACD68024;
XX
DT    17-SEP-2003 (first entry)
XX
DE   Novel human secreted and transmembrane protein PRO1303 cDNA.
XX
KW   Human; secreted and transmembrane protein; PRO; gene therapy; vaccine;
KW   tissue typing; chromosome identification; vaccine; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2003073129-A1.
XX
PD   17-APR-2003.
XX
FF   04-SEP-2001; 2001US-00946374.
XX
XX   01-SEP-1998; 98US-0098716P.
XX   01-SEP-1998; 98US-0098723P.
XX   01-SEP-1998; 98US-0098749P.
XX   01-SEP-1998; 98US-0098750P.
XX   02-SEP-1998; 98US-0098803P.
XX   02-SEP-1998; 98US-0098821P.
XX   02-SEP-1998; 98US-0098843P.
XX   09-SEP-1998; 98US-0099536P.
XX   09-SEP-1998; 98US-0099596P.
XX   09-SEP-1998; 98US-0099598P.
XX   09-SEP-1998; 98US-0099602P.
XX   10-SEP-1998; 98US-0099642P.
XX   10-SEP-1998; 98US-0099741P.
XX   10-SEP-1998; 98US-0099754P.
XX   10-SEP-1998; 98US-0099763P.
XX   10-SEP-1998; 98US-0099792P.

PR      10-SEP-1998; 98US-0098808P.
PR      10-SEP-1998; 98US-0098812P.
PR      10-SEP-1998; 98US-0098815P.
PR      10-SEP-1998; 98US-0098816P.
PR      15-SEP-1998; 98US-0100385P.
PR      15-SEP-1998; 98US-0100388P.
PR      15-SEP-1998; 98US-0100390P.
PR      16-SEP-1998; 98US-0100584P.
PR      16-SEP-1998; 98US-0100627P.
PR      16-SEP-1998; 98US-0100661P.
PR      16-SEP-1998; 98US-0100662P.
PR      16-SEP-1998; 98US-0100664P.
PR      17-SEP-1998; 98US-0100683P.
PR      17-SEP-1998; 98US-0100684P.
PR      17-SEP-1998; 98US-0100710P.
PR      17-SEP-1998; 98US-0100711P.
PR      17-SEP-1998; 98US-0100919P.
PR      17-SEP-1998; 98US-0100930P.
PR      18-SEP-1998; 98US-0100848P.
PR      18-SEP-1998; 98US-0100849P.
PR      18-SEP-1998; 98US-0101014P.
PR      18-SEP-1998; 98US-0101068P.
PR      18-SEP-1998; 98US-0101071P.
PR      22-SEP-1998; 98US-0101279P.
PR      23-SEP-1998; 98US-0101471P.
PR      23-SEP-1998; 98US-0101472P.
PR      23-SEP-1998; 98US-0101474P.
PR      23-SEP-1998; 98US-0101475P.
PR      23-SEP-1998; 98US-0101476P.
PR      23-SEP-1998; 98US-0101477P.
PR      24-SEP-1998; 98US-0101479P.
PR      24-SEP-1998; 98US-0101738P.
PR      24-SEP-1998; 98US-0101741P.
PR      24-SEP-1998; 98US-0101743P.
PR      24-SEP-1998; 98US-0101915P.
PR      24-SEP-1998; 98US-0101916P.
PR      29-SEP-1998; 98US-0102207P.
PR      29-SEP-1998; 98US-0102240P.
PR      29-SEP-1998; 98US-0102307P.
PR      29-SEP-1998; 98US-0102330P.
PR      30-SEP-1998; 98US-0102331P.
PR      30-SEP-1998; 98US-0102484P.
PR      30-SEP-1998; 98US-0102487P.
PR      30-SEP-1998; 98US-0102570P.
PR      30-SEP-1998; 98US-0102571P.
PR      01-OCT-1998; 98US-0102684P.
PR      02-OCT-1998; 98US-0102687P.
PR      06-OCT-1998; 98US-0102965P.
PR      06-OCT-1998; 98US-0103258P.
PR      07-OCT-1998; 98US-0103449P.
PR      07-OCT-1998; 98US-0103314P.
PR      07-OCT-1998; 98US-0103315P.
PR      07-OCT-1998; 98US-0103328P.
PR      07-OCT-1998; 98US-0103395P.
PR      07-OCT-1998; 98US-0103396P.
PR      08-OCT-1998; 98US-0103401P.
PR      08-OCT-1998; 98US-0103633P.
PR      08-OCT-1998; 98US-0103678P.
PR      08-OCT-1998; 98US-0103679P.
PR      08-OCT-1998; 98US-0103711P.
PR      14-OCT-1998; 98US-0104257P.
PR      20-OCT-1998; 98US-0104987P.
PR      20-OCT-1998; 98US-0105000P.
PR      20-OCT-1998; 98US-0105002P.
PR      21-OCT-1998; 98US-0105104P.
PR      22-OCT-1998; 98US-0105169P.
PR      22-OCT-1998; 98US-0105266P.
PR      26-OCT-1998; 98US-0105693P.
PR      26-OCT-1998; 98US-0105694P.
PR      27-OCT-1998; 98US-0105807P.
PR      27-OCT-1998; 98US-0105881P.
PR      27-OCT-1998; 98US-0105882P.
PR      27-OCT-1998; 98US-0106062P.
```

28-OCT-1998; 98US-0106023P.  
PR 28-OCT-1998; 98US-0106029P.  
PR 28-OCT-1998; 98US-0106030P.  
PR 28-OCT-1998; 98US-0106032P.  
PR 28-OCT-1998; 98US-0106033P.  
PR 28-OCT-1998; 98US-0106178P.  
PR 29-OCT-1998; 98US-0106248P.  
PR 29-OCT-1998; 98US-0106384P.  
PR 29-OCT-1998; 98US-0108500P.  
PR 30-OCT-1998; 98US-0108464P.  
PR 03-NOV-1998; 98US-0108856P.  
PR 03-NOV-1998; 98US-0106902P.  
PR 03-NOV-1998; 98US-0106905P.  
PR 03-NOV-1998; 98US-0106919P.  
PR 03-NOV-1998; 98US-0106932P.  
PR 03-NOV-1998; 98US-0106934P.  
PR 10-NOV-1998; 98US-0106934P.  
PR 17-NOV-1998; 98US-0107783P.  
PR 17-NOV-1998; 98US-0108775P.  
PR 17-NOV-1998; 98US-0108779P.  
PR 17-NOV-1998; 98US-0108787P.  
PR 17-NOV-1998; 98US-0108801P.  
PR 17-NOV-1998; 98US-0108802P.  
PR 17-NOV-1998; 98US-0108806P.  
PR 17-NOV-1998; 98US-0108807P.  
PR 17-NOV-1998; 98US-0108867P.  
PR 17-NOV-1998; 98US-0108925P.  
PR 18-NOV-1998; 98US-0108848P.  
PR 18-NOV-1998; 98US-0108849P.  
PR 18-NOV-1998; 98US-0108850P.  
PR 18-NOV-1998; 98US-0108851P.  
PR 18-NOV-1998; 98US-0108852P.  
PR 18-NOV-1998; 98US-0108858P.  
PR 18-NOV-1998; 98US-0108904P.  
PR 22-DEC-1998; 98US-00218317.  
PR 22-DEC-1998; 98US-0113296P.  
PR 30-DEC-1998; 98US-0114223P.  
PR 05-JAN-1999; 99WO-US000106.  
PR 12-APR-1999; 99US-00284291.  
PR 16-APR-1999; 99US-0129674P.  
PR 23-JUN-1999; 99US-0141037P.  
PR 20-JUL-1999; 99US-0144758P.  
PR 26-JUL-1999; 99US-0145698P.  
PR 01-SEP-1999; 99WO-US020111.  
PR 15-SEP-1999; 99WO-US021194.  
PR 18-OCT-1999; 99US-00403297.  
PR 30-NOV-1999; 99WO-US028313.  
PR 02-DEC-1999; 99WO-US028551.  
PR 16-DEC-1999; 99WO-US030095.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 15-MAR-2000; 2000WO-US005881.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 23-AUG-2000; 2000WO-US023524.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 01-JUN-2001; 2001WO-US00872035.  
PR 14-JUN-2001; 2001WO-US017800.  
PR 20-JUN-2001; 2001WO-US0082636.  
PR 29-JUN-2001; 2001WO-US019692.  
PR 09-JUL-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
  
(GETH ) GENENTECH INC.  
Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S,  
Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ,  
Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK,  
Williams PM, Wood WI,  
WPI; 2003-585292/55.  
DR P-PSDB; ABO33512.  
XX  
Novel isolated PRO polypeptides e.g. PRO1491 and PRO1571, useful in the  
preparation of a medicament for treating a condition responsive to PRO  
polypeptide, and as therapeutic agents e.g. vaccines.  
PS Claim 2; Fig 107; 561pp; English.  
XX  
The invention describes an isolated PRO (secreted and transmembrane)  
polypeptide (I), having at least 80% sequence identity to a sequence  
CC  
Query Match 100.0%; Score 1091; DB 8; Length 1091;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAAGCAGGTGATCCCTTGGTGGCTTCAAGAGAGACAGAGAGGCGAGAGTGGGGGGC 60  
DB 1 CAAGCAGGTGATCCCTTGGTGGCTTCAAGAGAGACAGAGAGGCGAGAGTGGGGGGC 60  
QY 61 ACAGGGAAGGGTGACCTCTGAGATCCCTTTTCCCCAGAGACTTTTGGAAAGTCAACCCACC 120  
DB 61 ACAGGGAAGGGTGACCTCTGAGATCCCTTTTCCCCAGAGACTTTTGGAAAGTCAACCCACC 120  
QY 121 ATGGGGCTCAGCATCTTTTGGTGGCTTGGTGGCTTGGGCTCAGCCAGGAGGAGCCG 180  
DB 121 ATGGGGCTCAGCATCTTTTGGTGGCTTGGTGGCTTGGGCTCAGCCAGGAGGAGCCG 180  
QY 181 AAGATTTTCAATGGCCTCAGTGTGGGCGTAACTCAGCCGCGGAGGAGTGGGGCTGTTT 240  
DB 181 AAGATTTTCAATGGCCTCAGTGTGGGCGTAACTCAGCCGCGGAGGAGTGGGGCTGTTT 240  
QY 241 GAGGGCACCAGCTGCGCTGGGGGGTGTCTTATTTGACACAGTGGGTCCTTCAAGCG 300  
DB 241 GAGGGCACCAGCTGCGCTGGGGGGTGTCTTATTTGACACAGTGGGTCCTTCAAGCG 300  
QY 301 GCTCACTGACGGGAGCAGAGTACTGGGTGGCGCTGGGGGAAACAGACCTCAGCAGCTC 360  
DB 301 GCTCACTGACGGGAGCAGAGTACTGGGTGGCGCTGGGGGAAACAGACCTCAGCAGCTC 360  
QY 361 GACTGACCGAGAGATCCGGCAGCAGCGGCTTCTGTGACCCATCCCGCTACCTGGGA 420  
DB 361 GACTGACCGAGAGATCCGGCAGCAGCGGCTTCTGTGACCCATCCCGCTACCTGGGA 420  
QY 421 GCTTGACGAGCAGCAGCAGCAGCTTGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTG 480  
DB 421 GCTTGACGAGCAGCAGCAGCAGCTTGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTG 480  
QY 481 ACCAGCAGGTTCAACCCCTGCGCCCTGCGCCATGACTGTGCAACCCATCCCGATCTGTCAG 600  
DB 481 ACCAGCAGGTTCAACCCCTGCGCCCTGCGCCATGACTGTGCAACCCATCCCGATCTGTCAG 600  
QY 601 TGCTTCAACCTCTCCATCGTCCCATGCGACCTTGGCTGGCTGGCTGGCTGGCTGGCTGG 660  
DB 601 TGCTTCAACCTCTCCATCGTCCCATGCGACCTTGGCTGGCTGGCTGGCTGGCTGGCTGG 660  
QY 661 ACAGGACACATGTTGTGTCAGGCGGCTCCCGGGGAGGATGCTGTCAGGAGTATCT 720  
DB 661 ACAGGACACATGTTGTGTCAGGCGGCTCCCGGGGAGGATGCTGTCAGGAGTATCT 720  
QY 721 GGGGGCCCCCTGGTGTGTGGGGAGTCTTCAAGTCTGTGTCTTGGGGTCTGTGGGG 780

```
Db      721 GGGGGCCCCCTGGTGTGGGGGAGTCTTCAAGGTCTGGTCTGGGGGCTGTGGGG 780
QY      781 CCTCTGGACAAGATGGCATCCTGGAGTCTACACCTATATTGCAAGTATGTGACTGG 840
Db      781 CCTCTGGACAAGATGGCATCCTGGAGTCTACACCTATATTGCAAGTATGTGACTGG 840
QY      841 ATCCGGATGATCATGAGGAACAACCTGACCTGTTCTCCACCTCCACCCCAACCCCTTAA 900
Db      841 ATCCGGATGATCATGAGGAACAACCTGACCTGTTCTCCACCTCCACCCCAACCCCTTAA 900
QY      901 CTTGGGTACCCCTCTGGCCCTCAGAGCAACCAATCTCTCATCATCTCCCTAGTCC 960
Db      901 CTTGGGTACCCCTCTGGCCCTCAGAGCAACCAATCTCTCATCATCTCCCTAGTCC 960
QY      961 ACTCTGTGGCTGGGAACTTCTTGAACCTTAACTCTGCCAGCCCTTCTTAAGACCA 1020
Db      961 ACTCTGTGGCTGGGAACTTCTTGAACCTTAACTCTGCCAGCCCTTCTTAAGACCA 1020
QY      1021 CGAGCGGGGTGAGAAAGTGTGCAATAGTCTGGAATAAATATAATGAAGGAGGGGCAAA 1080
Db      1021 CGAGCGGGGTGAGAAAGTGTGCAATAGTCTGGAATAAATATAATGAAGGAGGGGCAAA 1080
QY      1081 AAAAAAAAAA 1091
Db      1081 AAAAAAAAAA 1091

RESULT 8
ID      ADC18062
XX      ADC18062 standard; cDNA; 1091 BP.
AC      ADC18062;
XX      18-DEC-2003 (first entry)
XX      Human PRO polynucleotide #54.
KW      Human; PRO; gene; ss; protein electrophoresis; chromosome mapping;
KW      gene mapping; genetic disorder.
XX      Homo sapiens.
XX      US2003064925-A1.
XX      03-APR-2003.
XX      10-DEC-2001; 2001US-00013907.
XX      01-SEP-1998; 98US-0098716P.
XX      01-SEP-1998; 98US-0098723P.
XX      01-SEP-1998; 98US-0098749P.
XX      02-SEP-1998; 98US-0098750P.
XX      02-SEP-1998; 98US-0098803P.
XX      02-SEP-1998; 98US-0098821P.
XX      09-SEP-1998; 98US-0098843P.
XX      09-SEP-1998; 98US-0098936P.
XX      09-SEP-1998; 98US-0098966P.
XX      09-SEP-1998; 98US-0098988P.
XX      09-SEP-1998; 98US-0099602P.
XX      10-SEP-1998; 98US-0099642P.
XX      10-SEP-1998; 98US-0099741P.
XX      10-SEP-1998; 98US-0099754P.
XX      10-SEP-1998; 98US-0099763P.
XX      10-SEP-1998; 98US-0099792P.
XX      10-SEP-1998; 98US-0099808P.
XX      10-SEP-1998; 98US-0099812P.
XX      10-SEP-1998; 98US-0099815P.
XX      15-SEP-1998; 98US-0099816P.
XX      15-SEP-1998; 98US-0100385P.
XX      15-SEP-1998; 98US-0100388P.
XX      15-SEP-1998; 98US-0100390P.
XX      16-SEP-1998; 98US-0100584P.
XX      16-SEP-1998; 98US-0100627P.
XX      16-SEP-1998; 98US-0100661P.
XX      16-SEP-1998; 98US-0100662P.
XX      16-SEP-1998; 98US-0100664P.
XX      17-SEP-1998; 98US-0100683P.
XX      17-SEP-1998; 98US-0100684P.
XX      17-SEP-1998; 98US-0100710P.
XX      17-SEP-1998; 98US-0100711P.
XX      17-SEP-1998; 98US-0100919P.
XX      17-SEP-1998; 98US-0100930P.
XX      18-SEP-1998; 98US-0100848P.
XX      18-SEP-1998; 98US-0100849P.
XX      18-SEP-1998; 98US-0101014P.
XX      18-SEP-1998; 98US-0101068P.
XX      18-SEP-1998; 98US-0101071P.
XX      22-SEP-1998; 98US-0101279P.
XX      23-SEP-1998; 98US-0101471P.
XX      23-SEP-1998; 98US-0101472P.
XX      23-SEP-1998; 98US-0101474P.
XX      23-SEP-1998; 98US-0101475P.
XX      23-SEP-1998; 98US-0101476P.
XX      23-SEP-1998; 98US-0101477P.
XX      23-SEP-1998; 98US-0101479P.
XX      24-SEP-1998; 98US-0101738P.
XX      24-SEP-1998; 98US-0101741P.
XX      24-SEP-1998; 98US-0101743P.
XX      24-SEP-1998; 98US-0101915P.
XX      29-SEP-1998; 98US-0102070P.
XX      29-SEP-1998; 98US-0102240P.
XX      29-SEP-1998; 98US-0102307P.
XX      29-SEP-1998; 98US-0102330P.
XX      30-SEP-1998; 98US-0102331P.
XX      30-SEP-1998; 98US-0102484P.
XX      30-SEP-1998; 98US-0102570P.
XX      30-SEP-1998; 98US-0102571P.
XX      01-OCT-1998; 98US-0102684P.
XX      01-OCT-1998; 98US-0102877P.
XX      02-OCT-1998; 98US-0102965P.
XX      06-OCT-1998; 98US-0103258P.
XX      06-OCT-1998; 98US-0103449P.
XX      07-OCT-1998; 98US-0103314P.
XX      07-OCT-1998; 98US-0103315P.
XX      07-OCT-1998; 98US-0103288P.
XX      07-OCT-1998; 98US-0103395P.
XX      07-OCT-1998; 98US-0103396P.
XX      08-OCT-1998; 98US-0103401P.
XX      08-OCT-1998; 98US-0103633P.
XX      08-OCT-1998; 98US-0103678P.
XX      08-OCT-1998; 98US-0103679P.
XX      14-OCT-1998; 98US-0103711P.
XX      20-OCT-1998; 98US-0104257P.
XX      20-OCT-1998; 98US-0104987P.
XX      20-OCT-1998; 98US-0105000P.
XX      21-OCT-1998; 98US-0105002P.
XX      21-OCT-1998; 98US-0105104P.
XX      22-OCT-1998; 98US-0105169P.
XX      22-OCT-1998; 98US-0105266P.
XX      26-OCT-1998; 98US-0105693P.
XX      26-OCT-1998; 98US-0105694P.
XX      27-OCT-1998; 98US-0105807P.
XX      27-OCT-1998; 98US-0105881P.
XX      27-OCT-1998; 98US-0105882P.
XX      28-OCT-1998; 98US-0106023P.
XX      28-OCT-1998; 98US-0106029P.
XX      28-OCT-1998; 98US-0106030P.
XX      28-OCT-1998; 98US-0106032P.
XX      28-OCT-1998; 98US-0106033P.
XX      28-OCT-1998; 98US-0106178P.
XX      29-OCT-1998; 98US-0106248P.
XX      29-OCT-1998; 98US-0106384P.
```





```
Db      841 ATCCGATCATCATGAGAACACTGACCTGTTCTCCACCTCCACCCCTTAA 900
QY      901 CTTGGTACCTCTGCGCTCAGACACCAATATCTCTCCATCACTTCCCTAGCTCC 960
Db      901 CTTGGTACCTCTGCGCTCAGACACCAATATCTCTCCATCACTTCCCTAGCTCC 960
QY      961 ACTCTGTTGGCTGGGAATCTTCTGGAATTTAACTCTGCGAGCCCTTAAAGACCA 1020
Db      961 ACTCTGTTGGCTGGGAATCTTCTGGAATTTAACTCTGCGAGCCCTTAAAGACCA 1020
QY      1021 CGAGCGGTGAGAGAGTGTCAATAGTCTGGAATAAATAAATAAGAGGGGCAAA 1080
Db      1021 CGAGCGGTGAGAGAGTGTCAATAGTCTGGAATAAATAAATAAGAGGGGCAAA 1080
QY      1081 AAAAAAAAAA 1091
Db      1081 AAAAAAAAAA 1091

RESULT 9
ADD70708
ID ADD70708 standard; cDNA; 1091 bp.
XX
AC ADD70708;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human cDNA encoding secreted/transmembrane protein PRO1303.
XX
KW Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;
KW immune response; cardiac insufficiency disorder; calcium flux;
KW umbilical vein endothelial cell; bone disorder; cartilage disorder;
KW arthritis; wound healing; diabetes; skeletal muscle cells; obesity;
KW Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;
KW dermatitis; herpeticiformis; Crohn's disease; thalassaemia.
XX
OS Homo sapiens.
XX
US200309625-A1.
XX
29-MAY-2003.
XX
12-DEC-2001; 2001US-00015386.
XX
01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 01-SEP-1998; 98US-0098749P.
PR 01-SEP-1998; 98US-0098750P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 02-SEP-1998; 98US-0099536P.
PR 02-SEP-1998; 98US-0099596P.
PR 02-SEP-1998; 98US-0099598P.
PR 03-SEP-1998; 98US-0099602P.
PR 03-SEP-1998; 98US-0099642P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099792P.
PR 10-SEP-1998; 98US-0099808P.
PR 10-SEP-1998; 98US-0099812P.
PR 10-SEP-1998; 98US-0099815P.
PR 10-SEP-1998; 98US-0099816P.
PR 15-SEP-1998; 98US-0100385P.
PR 15-SEP-1998; 98US-0100388P.
PR 15-SEP-1998; 98US-0100390P.
PR 16-SEP-1998; 98US-0100584P.
PR 16-SEP-1998; 98US-0100627P.
PR 16-SEP-1998; 98US-0100661P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100710P.
PR 17-SEP-1998; 98US-0100711P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100848P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 18-SEP-1998; 98US-0101071P.
PR 22-SEP-1998; 98US-0101279P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101474P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101476P.
PR 23-SEP-1998; 98US-0101477P.
PR 23-SEP-1998; 98US-0101479P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101741P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101915P.
PR 24-SEP-1998; 98US-0101916P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102307P.
PR 29-SEP-1998; 98US-0102330P.
PR 30-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102484P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-0103314P.
PR 07-OCT-1998; 98US-0103315P.
PR 07-OCT-1998; 98US-0103328P.
PR 07-OCT-1998; 98US-0103395P.
PR 07-OCT-1998; 98US-0103396P.
PR 08-OCT-1998; 98US-0103401P.
PR 08-OCT-1998; 98US-0103633P.
PR 08-OCT-1998; 98US-0103678P.
PR 08-OCT-1998; 98US-0103679P.
PR 14-OCT-1998; 98US-0103711P.
PR 20-OCT-1998; 98US-0104257P.
PR 20-OCT-1998; 98US-0104987P.
PR 20-OCT-1998; 98US-0105002P.
PR 21-OCT-1998; 98US-0105104P.
PR 22-OCT-1998; 98US-0105169P.
PR 22-OCT-1998; 98US-0105266P.
PR 26-OCT-1998; 98US-0105693P.
PR 26-OCT-1998; 98US-0105694P.
PR 27-OCT-1998; 98US-0105807P.
PR 27-OCT-1998; 98US-0105881P.
PR 27-OCT-1998; 98US-0105882P.
PR 27-OCT-1998; 98US-0106062P.
PR 28-OCT-1998; 98US-0106023P.
PR 28-OCT-1998; 98US-0106029P.
PR 28-OCT-1998; 98US-0106030P.
PR 28-OCT-1998; 98US-0106032P.
PR 28-OCT-1998; 98US-0106178P.
PR 29-OCT-1998; 98US-0106248P.
PR 29-OCT-1998; 98US-0106384P.
PR 29-OCT-1998; 98US-0108500P.
PR 30-OCT-1998; 98US-0106464P.
PR 03-NOV-1998; 98US-0106856P.
PR 03-NOV-1998; 98US-0106902P.
```



Db 961 ACTCTGTTGGCTGGGAATCTTCTGGAACCTTAACTCTCGCAGCCCTTCTAAGACCCA 1020  
 QY 1021 CGAGCGGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATATAAATGAAGAGGGGCAAA 1080  
 Db 1021 CGAGCGGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATATAAATGAAGAGGGGCAAA 1080  
 QY 1081 AAAAAAAAAA 1091  
 Db 1081 AAAAAAAAAA 1091

RESULT 10  
 ADD39785  
 ID ADD39785 standard; cDNA; 1091 BP.  
 AC ADD39785;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Human cDNA encoding secreted/transmembrane protein PRO1303.  
 XX  
 KW Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;  
 KW immune response; cardiac insufficiency disorder; calcium flux;  
 KW umbilical vein endothelial cell; bone disorder; cartilage disorder;  
 KW arthritis; wound healing; diabetes; skeletal muscle cells; obesity;  
 KW Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;  
 KW dermatitis; herpetiformis; Crohn's disease; thalassaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003083462-Al.  
 XX  
 PD 01-MAY-2003.  
 XX  
 PF 10-DEC-2001; 2001US-00013913.  
 XX  
 PR 05-JAN-1999; 99WO-US000106.  
 PR 01-SEP-1999; 99WO-US020111.  
 PR 15-SEP-1999; 99WO-US021134.  
 PR 30-NOV-1999; 99WO-US028313.  
 PR 02-DEC-1999; 99WO-US028551.  
 PR 16-DEC-1999; 99WO-US030095.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000376.  
 PR 11-FEB-2000; 2000WO-US003565.  
 PR 18-FEB-2000; 2000WO-US004342.  
 PR 24-FEB-2000; 2000WO-US005004.  
 PR 02-MAR-2000; 2000WO-US005841.  
 PR 15-MAR-2000; 2000WO-US006884.  
 PR 17-MAY-2000; 2000WO-US013705.  
 PR 22-MAY-2000; 2000WO-US014042.  
 PR 30-MAY-2000; 2000WO-US014941.  
 PR 02-JUN-2000; 2000WO-US015264.  
 PR 23-AUG-2000; 2000WO-US023522.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 08-NOV-2000; 2000WO-US030952.  
 PR 10-NOV-2000; 2000WO-US030873.  
 PR 01-DEC-2000; 2000WO-US032678.  
 PR 28-FEB-2001; 2001WO-US006520.  
 PR 01-MAR-2001; 2001WO-US006666.  
 PR 01-JUN-2001; 2001WO-US017800.  
 PR 20-JUN-2001; 2001WO-US019692.  
 PR 29-JUN-2001; 2001WO-US021066.  
 PR 09-JUL-2001; 2001WO-US021735.  
 PR 04-SEP-2001; 2001US-00946374.  
 XX  
 PA (GETH ) GENENTECH INC.

Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;  
 Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;  
 Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe  
 Williams PM, Wood WI;

XX WPI; 2003-755122/71.  
 DR P-PSDB; ADD39786.  
 DR  
 XX  
 PT New secreted and transmembrane PRO polypeptides useful for treating  
 PT cancers, kidney disorders, Crohn's disease, diabetes mellitus, hyper- or  
 PT hypo-insulinemia, sports injuries and arthritis.  
 XX  
 PS Claim 2; SEQ ID NO 193; 557pp; English.  
 XX  
 CC The invention relates to an isolated PRO polypeptide (secreted or  
 CC transmembrane protein) having at least 80% amino acid sequence identity  
 CC to an amino acid sequence chosen from 123 fully defined sequences as  
 CC given in the specification (including their extracellular domains either  
 CC or without their associated signal peptides. Also include are the  
 CC nucleotide (NA) sequences encoding PRO, a vector comprising the PRO NA, a  
 CC host cell comprising the vector, producing PRO, a chimeric molecule  
 CC comprising PRO fused to a heterologous amino acid sequence, and an anti-  
 CC PRO antibody. PRO is useful as molecular weight markers for protein  
 CC electrophoresis and also for chromosome identification. PRO is also  
 CC useful for tissue typing. PRO and PRO NA are useful as hybridisation  
 CC probes for a cDNA library to isolate the full-length PRO cDNA. PRO NA is  
 CC useful for generating transgenic animals or knock-out animals which are  
 CC useful in development and screening useful reagents. PRO NA is also  
 CC useful in gene therapy. PRO1244, PRO1286 and PRO1303 polypeptides are  
 CC useful for treating cancerous tumours. PRO1250, PRO1418 and PRO1410  
 CC polypeptides are useful for suppressing immune response. PRO1246  
 CC polypeptide is useful for treating cardiac insufficiency disorders.  
 CC PRO1246 polypeptide is also useful for treating tumours. PRO1246 and  
 CC PRO1561 polypeptide are useful for stimulating calcium flux in human  
 CC umbilical vein endothelial cells. PRO1265, PRO1250 and PRO1474  
 CC polypeptides are useful for treating bone and/or cartilage disorders  
 CC (e.g., arthritis) and wound healing. PRO1130, PRO1275 and PRO1418  
 CC polypeptides are useful for treating diabetes in skeletal muscle cells  
 CC and obesity. PRO1265, PRO1244 and PRO1382 polypeptides are useful for  
 CC treating Berger disease or other nephropathies associated with Schonlein-  
 CC disease. PRO1478, PRO1265, PRO1412, PRO1279, PRO1304, PRO1306, PRO1418,  
 CC PRO1410 and PRO1575 are useful in treating thalassaemias. The present  
 CC sequence encodes a PRO protein of the invention.  
 XX  
 SQ Sequence 1091 BP; 221 A; 335 C; 313 G; 222 T; 0 U; 0 Other;

Query Match 100.0%; Score 1091; DB 9; Length 1091;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAAGCAGGTCTATCCCTTGGTGCCTTCAAGAGAGCAGAGGGGAGAGTGGGGGGC 60  
 Db 1 CAAGCAGGTCTATCCCTTGGTGCCTTCAAGAGAGCAGAGGGGAGAGTGGGGGGC 60  
 QY 61 ACAGGGAAGAGGTGACCTCTGAGATTCCTTTTCCCGCAGACTTTGGAGTGCACCCACC 120  
 Db 61 ACAGGGAAGAGGTGACCTCTGAGATTCCTTTTCCCGCAGACTTTGGAGTGCACCCACC 120  
 QY 121 ATGGGGCTCAGCATCTTTTGGTCTCTGTGTGTTTGGGCTCAGCCAGGAGCACCAG 180  
 Db 121 ATGGGGCTCAGCATCTTTTGGTCTCTGTGTGTTTGGGCTCAGCCAGGAGCACCAG 180  
 QY 181 AAGATTTCATATGGCACTGAGTGTGGGCGTAACTCAGCCGCTGGGAGTGGGGCTGTTT 240  
 Db 181 AAGATTTCATATGGCACTGAGTGTGGGCGTAACTCAGCCGCTGGGAGTGGGGCTGTTT 240  
 QY 241 GAGGGCACCAGCTGGCTGCGGGGTGTCCTTATTGACACAGAGTGGGTCTCTCAGAG 300  
 Db 241 GAGGGCACCAGCTGGCTGCGGGGTGTCCTTATTGACACAGAGTGGGTCTCTCAGAG 300  
 QY 301 GCTCAGTCAGCGGAGCAGGTACTGGGTGCGCTGGGGGAAACACAGCTCAGCCAGCTC 360  
 Db 301 GCTCAGTCAGCGGAGCAGGTACTGGGTGCGCTGGGGGAAACACAGCTCAGCCAGCTC 360  
 QY 361 GACTGACCCAGCAGATCCGACAGCGGCTTCTCTGTGACCATCCCGGCTACCTGGGA 420

Db 361 GACTGGACCGAGCAGATCCGGCAGCAGCGCTTCTGTGTGACCCATCCCGGCTACCTGGGA 420  
Qy 421 GCCTCGAGGACCGAGCAGCAGCTCCGGTGTCTGGGCTGGGCTGCCGTCCCGTCCGGGTA 480  
Db 421 GCCTCGAGGACCGAGCAGCAGCTCCGGTGTCTGGGCTGGGCTGCCGTCCCGTCCGGGTA 480  
Qy 481 ACCAGCAGCGTTCAACCCCTGCCCTGCCCAATGACTGTGCAACCGCTGGCACCGAGTGC 540  
Db 481 ACCAGCAGCGTTCAACCCCTGCCCTGCCCAATGACTGTGCAACCGCTGGCACCGAGTGC 540  
Qy 541 CACGTCTCAGCTGGGGATCACCAACACCGAACCCATTCGCCGATCTCTCCAG 600  
Db 541 CACGTCTCAGCTGGGGATCACCAACACCGAACCCATTCGCCGATCTCTCCAG 600  
Qy 601 TGCCTCAACCTCTCCATGCTCTCCATGCCACCTGCCATGCTGTATCCCGGAGATC 660  
Db 601 TGCCTCAACCTCTCCATGCTCTCCATGCCACCTGCCATGCTGTATCCCGGAGATC 660  
Qy 661 ACAGCAACATGTGTGTGAGGCGGGTCTCCGGGGCAGGATGCTGCCAGGGTGTCT 720  
Db 661 ACAGCAACATGTGTGTGAGGCGGGTCTCCGGGGCAGGATGCTGCCAGGGTGTCT 720  
Qy 721 GGGGGCCCCCTGTGTGTGGGGAGTCTTCAAGTCTGTGTCTCCGGGGTCTGTGGGG 780  
Db 721 GGGGGCCCCCTGTGTGTGGGGAGTCTTCAAGTCTGTGTCTCCGGGGTCTGTGGGG 780  
Qy 781 CCCTGTGGACAAGATGCATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGGACTGG 840  
Db 781 CCCTGTGGACAAGATGCATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGGACTGG 840  
Qy 841 ATCCGATGATCATGAGGAACAACCTGACCTGTCTTCCCTCCACCTCCACCCCAACCCCTTAA 900  
Db 841 ATCCGATGATCATGAGGAACAACCTGACCTGTCTTCCCTCCACCTCCACCCCAACCCCTTAA 900  
Qy 901 CTTGGGTACCCCTCTGGCCCTCAGAGCAACATATCTCTCCATCTCCCTAGCTCC 960  
Db 901 CTTGGGTACCCCTCTGGCCCTCAGAGCAACATATCTCTCCATCTCCCTAGCTCC 960  
Qy 961 ACTCTTTGTCCTGGGAACTCTTGAACTTTAACTCTCCGAGCCCTCTTAAGACCCA 1020  
Db 961 ACTCTTTGTCCTGGGAACTCTTGAACTTTAACTCTCCGAGCCCTCTTAAGACCCA 1020  
Qy 1021 CGAGCGGGTGAGAGAGTGTCAATAGTGTGAATAATAATAATAAGAGGGGCAAA 1080  
Db 1021 CGAGCGGGTGAGAGAGTGTCAATAGTGTGAATAATAATAATAAGAGGGGCAAA 1080  
Qy 1081 AAAAAAAAAA 1091  
Db 1081 AAAAAAAAAA 1091

RESULT 11

ADD70231  
ID ADD70231 standard; cDNA; 1091 BP.  
XX  
AC ADD70231;  
DT  
DT 15-JAN-2004 (first entry)  
XX  
DE Human cDNA encoding secreted/transmembrane protein PRO1303.  
XX  
KW Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;  
KW immune response; cardiac insufficiency disorder; calcium flux;  
KW umbilical vein endothelial cell; bone disorder; cartilage disorder;  
KW arthritis; wound healing; diabetes; skeletal muscle cells; obesity;  
KW Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;  
KW dermatitis; herpeticiformis; Crohn's disease; thalassaemia.  
XX  
OS Homo sapiens.  
XX  
FN US2003054406-A1.  
XX  
PD 20-MAR-2003.

XX 06-DEC-2001; 2001US-00006818.  
PF 01-SEP-1998; 98US-0098716P.  
XX 01-SEP-1998; 98US-0098723P.  
PR 01-SEP-1998; 98US-0098749P.  
PR 01-SEP-1998; 98US-0098750P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099536P.  
PR 09-SEP-1998; 98US-0099598P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 09-SEP-1998; 98US-0099642P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099792P.  
PR 10-SEP-1998; 98US-0099808P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 10-SEP-1998; 98US-0099815P.  
PR 10-SEP-1998; 98US-0099816P.  
PR 15-SEP-1998; 98US-0100385P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 15-SEP-1998; 98US-0100390P.  
PR 16-SEP-1998; 98US-0100584P.  
PR 16-SEP-1998; 98US-0100627P.  
PR 16-SEP-1998; 98US-0100661P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100710P.  
PR 17-SEP-1998; 98US-0100711P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100848P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 18-SEP-1998; 98US-0101071P.  
PR 22-SEP-1998; 98US-0101279P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101474P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101476P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101479P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101741P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101915P.  
PR 24-SEP-1998; 98US-0101916P.  
PR 29-SEP-1998; 98US-010207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102307P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 30-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102484P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 01-OCT-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 02-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-0103314P.  
PR 07-OCT-1998; 98US-0103315P.  
PR 07-OCT-1998; 98US-0103328P.  
PR 07-OCT-1998; 98US-0103395P.









```
Db      601  TGCCTCAACCTCTCCATCGTCTCCCATGCCACCTGCCATGGTGTGTATCCCGGAGAAATC 660
Qy      661  ACGAGCAACATGGTGTGTGCGAGGCGCTCCCGGCGCAGGATGCTGCCAGGGTGATTCT 720
Db      661  ACGAGCAACATGGTGTGTGCGAGGCGCTCCCGGCGCAGGATGCTGCCAGGGTGATTCT 720
Qy      721  GGGGGCCCCCTGGTGTGTGCGGGGAGTCCTTCAAGSTCTGGTGTCTCTGGGGGTCTGTGGG 780
Db      721  GGGGGCCCCCTGGTGTGTGCGGGGAGTCCTTCAAGSTCTGGTGTCTCTGGGGGTCTGTGGG 780
Qy      781  CCCTGTGACAAAGATGGCATCCCTGCAGTCTACACCTATATTTGCAAGTATGTGACTGG 840
Db      781  CCCTGTGACAAAGATGGCATCCCTGCAGTCTACACCTATATTTGCAAGTATGTGACTGG 840
Qy      841  ATCCGGATGATCATGAGAAACAATGACCTGTTTCTCCACCTCCACCCCAACCCCTTAA 900
Db      841  ATCCGGATGATCATGAGAAACAATGACCTGTTTCTCCACCTCCACCCCAACCCCTTAA 900
Qy      901  CTTGGGTACCCCTCGGCCCTCAGAGCACCATACTCTCCATCATTCCCTAGCTCC 960
Db      901  CTTGGGTACCCCTCGGCCCTCAGAGCACCATACTCTCCATCATTCCCTAGCTCC 960
Qy      961  ACTCTGTGTGGCTCGGAACTTCTTGGAACTTTAACTCTGCGAGCCCTTCTAAGACCCA 1020
Db      961  ACTCTGTGTGGCTCGGAACTTCTTGGAACTTTAACTCTGCGAGCCCTTCTAAGACCCA 1020
Qy      1021  CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAAATAAATAAATGAAGGGGGCAAA 1080
Db      1021  CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAAATAAATAAATGAAGGGGGCAAA 1080
Qy      1081  AAAAAAAAAA 1091
Db      1081  AAAAAAAAAA 1091

RESULT 13
ADD33308
ID  ADD33308 standard; cDNA; 1091 BP.
AC  ADD33308;
XX
XX
DT  15-JAN-2004 (first entry)
XX
DE  Human cDNA encoding secreted/transmembrane protein PRO1303.
XX
KW  Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;
KW  immune response; cardiac insufficiency disorder; calcium flux;
KW  umbilical vein endothelial cell; bone disorder; cartilage disorder;
KW  arthritis; wound healing; diabetes; skeletal muscle cells; obesity;
KW  Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;
KW  dermatitis; herpeticiformis; Crohn's disease; thalassaemia.
XX
OS  Homo sapiens.
XX
XX  US2003096954-A1.
PN  22-MAY-2003.
PD
XX
XX  07-DEC-2001; 2001US-00011671.
PF
XX
XX  01-SEP-1998; 98US-0098716P.
PR  01-SEP-1998; 98US-0098723P.
PR  01-SEP-1998; 98US-0098749P.
PR  01-SEP-1998; 98US-0098750P.
PR  02-SEP-1998; 98US-0098803P.
PR  02-SEP-1998; 98US-0098821P.
PR  02-SEP-1998; 98US-0098843P.
PR  09-SEP-1998; 98US-0099536P.
PR  09-SEP-1998; 98US-0099598P.
PR  09-SEP-1998; 98US-0099602P.
PR  09-SEP-1998; 98US-0099642P.
PR  10-SEP-1998; 98US-0099741P.
PR  10-SEP-1998; 98US-0099754P.
PR  10-SEP-1998; 98US-0099763P.
PR  10-SEP-1998; 98US-0099792P.
PR  10-SEP-1998; 98US-0099808P.
PR  10-SEP-1998; 98US-0099812P.
PR  10-SEP-1998; 98US-0099815P.
PR  10-SEP-1998; 98US-0099816P.
PR  15-SEP-1998; 98US-0100385P.
PR  15-SEP-1998; 98US-0100388P.
PR  15-SEP-1998; 98US-0100390P.
PR  16-SEP-1998; 98US-0100584P.
PR  16-SEP-1998; 98US-0100627P.
PR  16-SEP-1998; 98US-0100661P.
PR  16-SEP-1998; 98US-0100662P.
PR  16-SEP-1998; 98US-0100664P.
PR  17-SEP-1998; 98US-0100683P.
PR  17-SEP-1998; 98US-0100684P.
PR  17-SEP-1998; 98US-0100710P.
PR  17-SEP-1998; 98US-0100711P.
PR  17-SEP-1998; 98US-0100919P.
PR  17-SEP-1998; 98US-0100930P.
PR  18-SEP-1998; 98US-0100848P.
PR  18-SEP-1998; 98US-0100849P.
PR  18-SEP-1998; 98US-0101014P.
PR  18-SEP-1998; 98US-0101068P.
PR  18-SEP-1998; 98US-0101071P.
PR  22-SEP-1998; 98US-0101279P.
PR  22-SEP-1998; 98US-0101471P.
PR  23-SEP-1998; 98US-0101472P.
PR  23-SEP-1998; 98US-0101474P.
PR  23-SEP-1998; 98US-0101475P.
PR  23-SEP-1998; 98US-0101476P.
PR  23-SEP-1998; 98US-0101477P.
PR  23-SEP-1998; 98US-0101479P.
PR  24-SEP-1998; 98US-0101738P.
PR  24-SEP-1998; 98US-0101741P.
PR  24-SEP-1998; 98US-0101743P.
PR  24-SEP-1998; 98US-0101916P.
PR  29-SEP-1998; 98US-0102207P.
PR  29-SEP-1998; 98US-0102240P.
PR  29-SEP-1998; 98US-0102307P.
PR  29-SEP-1998; 98US-0102330P.
PR  29-SEP-1998; 98US-0102331P.
PR  30-SEP-1998; 98US-0102484P.
PR  30-SEP-1998; 98US-0102487P.
PR  30-SEP-1998; 98US-0102570P.
PR  30-SEP-1998; 98US-0102571P.
PR  01-OCT-1998; 98US-0102684P.
PR  01-OCT-1998; 98US-0102687P.
PR  02-OCT-1998; 98US-0102965P.
PR  06-OCT-1998; 98US-0103258P.
PR  06-OCT-1998; 98US-0103449P.
PR  07-OCT-1998; 98US-0103314P.
PR  07-OCT-1998; 98US-0103315P.
PR  07-OCT-1998; 98US-0103328P.
PR  07-OCT-1998; 98US-0103395P.
PR  07-OCT-1998; 98US-0103396P.
PR  07-OCT-1998; 98US-0103401P.
PR  08-OCT-1998; 98US-0103633P.
PR  08-OCT-1998; 98US-0103678P.
PR  08-OCT-1998; 98US-0103679P.
PR  08-OCT-1998; 98US-0103711P.
PR  14-OCT-1998; 98US-0104257P.
PR  20-OCT-1998; 98US-0104987P.
PR  20-OCT-1998; 98US-0105000P.
PR  20-OCT-1998; 98US-0105002P.
PR  21-OCT-1998; 98US-0105104P.
PR  22-OCT-1998; 98US-0105169P.
PR  22-OCT-1998; 98US-0105266P.
PR  26-OCT-1998; 98US-0105693P.
PR  26-OCT-1998; 98US-0105694P.
PR  27-OCT-1998; 98US-0105807P.
```



Db 721 GGGGGCCCCCTGGTGTGGGGGAGTCCTTCAAGGTCTGGTGTCTCTGGGGCTCTGGGG 780  
Qy 781 CCCTGTGGAAGAATGGCATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGACATGG 840  
Db 781 CCCTGTGGAAGAATGGCATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGACATGG 840  
Qy 841 ATCCGGATGATCATGAGGAACAACCTGACCTGTTTCTCCACCTCCACCCCAACCCCTTAA 900  
Db 841 ATCCGGATGATCATGAGGAACAACCTGACCTGTTTCTCCACCTCCACCCCAACCCCTTAA 900  
Qy 901 CTTGGGTACCCCTCTGGGCCCTCAGAGCAACCAATATCTCTCCATCACCCTCCCTAGTCC 960  
Db 901 CTTGGGTACCCCTCTGGGCCCTCAGAGCAACCAATATCTCTCCATCACCCTCCCTAGTCC 960  
Qy 961 ACTCTTGTGGCTGGGAACCTTCTTGAACCTTTAACTCTCTGCGAGCCCTTCTAAGACCA 1020  
Db 961 ACTCTTGTGGCTGGGAACCTTCTTGAACCTTTAACTCTCTGCGAGCCCTTCTAAGACCA 1020  
Qy 1021 CGAGCGGGTGCAGAACTGTGCAATAGTCTGGAATAATAATGAAGGGGCAAA 1080  
Db 1021 CGAGCGGGTGCAGAACTGTGCAATAGTCTGGAATAATAATGAAGGGGCAAA 1080  
Qy 1081 AAAAAAAAAA 1091  
Db 1081 AAAAAAAAAA 1091  
  
RESULT 14  
ADD38831  
ID ADD38831 standard; cDNA; 1091 BP.  
XX  
AC ADD38831;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Human cDNA encoding secreted/transmembrane protein PRO1303.  
XX  
KW Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;  
KW immune response; cardiac insufficiency disorder; calcium flux;  
KW umbilical vein endothelial cell; bone disorder; cartilage disorder;  
KW arthritis; wound healing; diabetes; skeletal muscle cells; obesity;  
KW Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;  
KW dermatitis; herpeticiformis; Crohn's disease; thalassaemia.  
XX  
OS Homo sapiens.  
XX  
FN US2003092061-A1.  
XX  
PD 15-MAY-2003.  
XX  
PF 06-DEC-2001; 2001US-00007194.  
XX  
PR 01-SEP-1998; 98US-0038716P.  
PR 01-SEP-1998; 98US-0038723P.  
PR 01-SEP-1998; 98US-0038749P.  
PR 01-SEP-1998; 98US-0038750P.  
PR 02-SEP-1998; 98US-0038803P.  
PR 02-SEP-1998; 98US-0038821P.  
PR 02-SEP-1998; 98US-0038843P.  
PR 09-SEP-1998; 98US-0039536P.  
PR 09-SEP-1998; 98US-0039596P.  
PR 09-SEP-1998; 98US-0039602P.  
PR 09-SEP-1998; 98US-0039642P.  
PR 10-SEP-1998; 98US-0039741P.  
PR 10-SEP-1998; 98US-0039754P.  
PR 10-SEP-1998; 98US-0039763P.  
PR 10-SEP-1998; 98US-0039792P.  
PR 10-SEP-1998; 98US-0039808P.  
PR 10-SEP-1998; 98US-0039812P.  
PR 10-SEP-1998; 98US-0039815P.  
PR 10-SEP-1998; 98US-0039816P.  
PR 15-SEP-1998; 98US-0100385P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 15-SEP-1998; 98US-0100390P.  
PR 16-SEP-1998; 98US-0100584P.  
PR 16-SEP-1998; 98US-0100627P.  
PR 16-SEP-1998; 98US-0100661P.  
PR 16-SEP-1998; 98US-0100682P.  
PR 16-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100710P.  
PR 17-SEP-1998; 98US-0100711P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100848P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 18-SEP-1998; 98US-0101071P.  
PR 22-SEP-1998; 98US-0101279P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101474P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101476P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 23-SEP-1998; 98US-0101479P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101741P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101915P.  
PR 24-SEP-1998; 98US-0101916P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102307P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 30-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102484P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 01-OCT-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 02-OCT-1998; 98US-0102687P.  
PR 06-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 07-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-0103314P.  
PR 07-OCT-1998; 98US-0103315P.  
PR 07-OCT-1998; 98US-0103328P.  
PR 07-OCT-1998; 98US-0103395P.  
PR 07-OCT-1998; 98US-0103396P.  
PR 07-OCT-1998; 98US-0103401P.  
PR 08-OCT-1998; 98US-0103633P.  
PR 08-OCT-1998; 98US-0103678P.  
PR 08-OCT-1998; 98US-0103679P.  
PR 08-OCT-1998; 98US-0103711P.  
PR 14-OCT-1998; 98US-0104257P.  
PR 20-OCT-1998; 98US-0104987P.  
PR 20-OCT-1998; 98US-0105000P.  
PR 21-OCT-1998; 98US-0105002P.  
PR 21-OCT-1998; 98US-0105104P.  
PR 22-OCT-1998; 98US-0105169P.  
PR 22-OCT-1998; 98US-0105266P.  
PR 26-OCT-1998; 98US-0105693P.  
PR 26-OCT-1998; 98US-0105694P.  
PR 27-OCT-1998; 98US-0105807P.  
PR 27-OCT-1998; 98US-0105881P.  
PR 27-OCT-1998; 98US-0105882P.  
PR 28-OCT-1998; 98US-0106023P.  
PR 28-OCT-1998; 98US-0106029P.  
PR 28-OCT-1998; 98US-0106030P.  
PR 28-OCT-1998; 98US-0106032P.  
PR 28-OCT-1998; 98US-0106033P.



Db 841 ATCCGATGATCATGAGGACAACTGACCTGTTTCTCCACCTCCACCCACCCCTTAA 900  
QY 901 CTTGGGTACCCCTCGGCCCTCAGAGACCAATATCTCTCCATCACTTCCCTAGCTCC 960  
Db 901 CTTGGGTACCCCTCGGCCCTCAGAGACCAATATCTCTCCATCACTTCCCTAGCTCC 960  
QY 961 ACTCTGTTGGCTGGGAACTCTTGGAACTTTAACTCTGCGAGCCCTTCTAAGACCCA 1020  
Db 961 ACTCTGTTGGCTGGGAACTCTTGGAACTTTAACTCTGCGAGCCCTTCTAAGACCCA 1020  
QY 1021 CGAGCGGGTGAGAGAGTGTCAATGATGCTGGAATAAATAATGAAGAGGGGCAAA 1080  
Db 1021 CGAGCGGGTGAGAGAGTGTCAATGATGCTGGAATAAATAATGAAGAGGGGCAAA 1080  
QY 1081 AAAAAAAAAA 1091  
Db 1081 AAAAAAAAAA 1091  
RESULT 15  
ADD40262  
ID ADD40262 standard; cDNA; 1091 BP.  
AC ADD40262;  
XX  
XX  
DT 15-JAN-2004 (first entry)  
DE Human cDNA encoding secreted/transmembrane protein PRO1303.  
XX Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;  
KW immune response; cardiac insufficiency disorder; calcium flux;  
KW umbilical vein endothelial cell; bone disorder; cartilage disorder;  
KW arthritis; wound healing; diabetes; skeletal muscle cells; obesity;  
KW Berger disease; nephropathy; Schonlein-Henoch purpura; colliac disease;  
KW Dermatitis; herpeticiformis; Crohn's disease; thalassaemia.  
XX  
XX Homo sapiens.  
XX  
XX US2003082627-A1.  
PN  
XX  
PD 01-MAY-2003.  
XX  
XX 06-DEC-2001; 2001US-00006117.  
PF  
PF 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 01-SEP-1998; 98US-0098749P.  
PR 01-SEP-1998; 98US-0098750P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099536P.  
PR 09-SEP-1998; 98US-0099596P.  
PR 09-SEP-1998; 98US-0099596P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 09-SEP-1998; 98US-0099642P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099792P.  
PR 10-SEP-1998; 98US-0099808P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 10-SEP-1998; 98US-0099815P.  
PR 10-SEP-1998; 98US-0099816P.  
PR 15-SEP-1998; 98US-0100385P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 15-SEP-1998; 98US-0100390P.  
PR 16-SEP-1998; 98US-0100584P.  
PR 16-SEP-1998; 98US-0100627P.  
PR 16-SEP-1998; 98US-0100661P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 17-SEP-1998; 98US-0100663P.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100710P.  
PR 17-SEP-1998; 98US-0100711P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100848P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 18-SEP-1998; 98US-0101071P.  
PR 22-SEP-1998; 98US-0101279P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101474P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101476P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 23-SEP-1998; 98US-0101479P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101741P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101915P.  
PR 24-SEP-1998; 98US-0101916P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102307P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102484P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-0103314P.  
PR 07-OCT-1998; 98US-0103315P.  
PR 07-OCT-1998; 98US-0103328P.  
PR 07-OCT-1998; 98US-0103395P.  
PR 07-OCT-1998; 98US-0103396P.  
PR 07-OCT-1998; 98US-0103401P.  
PR 08-OCT-1998; 98US-0103633P.  
PR 08-OCT-1998; 98US-0103678P.  
PR 08-OCT-1998; 98US-0103679P.  
PR 08-OCT-1998; 98US-0103711P.  
PR 20-OCT-1998; 98US-0104257P.  
PR 20-OCT-1998; 98US-0104987P.  
PR 20-OCT-1998; 98US-0105000P.  
PR 20-OCT-1998; 98US-0105002P.  
PR 21-OCT-1998; 98US-0105104P.  
PR 22-OCT-1998; 98US-0105169P.  
PR 22-OCT-1998; 98US-0105266P.  
PR 26-OCT-1998; 98US-0105693P.  
PR 26-OCT-1998; 98US-0105694P.  
PR 27-OCT-1998; 98US-0105807P.  
PR 27-OCT-1998; 98US-0105881P.  
PR 27-OCT-1998; 98US-0105882P.  
PR 27-OCT-1998; 98US-0106023P.  
PR 28-OCT-1998; 98US-0106029P.  
PR 28-OCT-1998; 98US-0106030P.  
PR 28-OCT-1998; 98US-0106032P.  
PR 28-OCT-1998; 98US-0106033P.  
PR 28-OCT-1998; 98US-0106178P.  
PR 29-OCT-1998; 98US-0106248P.  
PR 29-OCT-1998; 98US-0106384P.  
PR 29-OCT-1998; 98US-0108500P.  
PR 30-OCT-1998; 98US-0108646P.  
PR 03-NOV-1998; 98US-0108656P.  
PR 03-NOV-1998; 98US-0106902P.  
PR 03-NOV-1998; 98US-0106905P.



Db 961 ACTCTTTGTCCTGGGNACTTCTTGGNACTTTAACTCTCTGCCAGCCCTTCTTAGACCCA 1020  
QY 1021 CGAGCGGGTGAGAGAACTGTGCAATAGTCTGGAAATAAATAAATGAAGAGGGGCAAA 1080  
Db 1021 CGAGCGGGTGAGAGAACTGTGCAATAGTCTGGAAATAAATAAATGAAGAGGGGCAAA 1080  
QY 1081 AAAAAAAAAA 1091  
Db 1081 AAAAAAAAAA 1091

Search completed: June 20, 2004, 06:17:42  
Job time : 512 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 06:09:07 ; Search time 3353 Seconds

(without alignments)  
9716.564 Million cell updates/sec

Title: US-10-015-385A-193

Perfect score: 1091

Sequence: 1 caagcaggtcatcccttgg.....aggggcaaaaaaaaaa 1091

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

```
EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmd.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_nam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	325	29.8	472	13	BX110362
2	254	23.3	469	9	AL394679
3	250	22.9	1186	11	BC035385
4	216	19.8	264	12	BG230967

5	101	9.3	383	9	AI554072
c	6	9.3	478	13	BX103516
7	52	4.8	550	12	BG721150
c	8	2.4	577	13	BQ129094
9	25	2.3	600	13	BQ270082
10	25	2.3	636	13	BQ269009
c	11	2.2	283	12	BQ794700
12	23	2.1	449	9	AI322408
c	13	2.1	451	14	CA313295
14	23	2.1	506	9	AA028356
15	23	2.1	520	13	BX514743
c	16	2.1	545	9	AI324852
17	23	2.1	627	29	CE842883
18	23	2.1	763	14	CD646427
19	23	2.1	765	12	BG762809
20	23	2.1	768	29	CC558299
21	23	2.1	890	11	AK009217
22	23	2.1	898	13	BY709025
c	23	2.1	904	13	EX689949
c	24	2.1	1002	29	CNS05FCB
c	25	2.0	149	12	BI278918
c	26	2.0	224	9	AI924035
c	27	2.0	242	10	BF353225
c	28	2.0	255	29	CE737502
c	29	2.0	287	12	BM368334
c	30	2.0	351	10	BF073014
31	22	2.0	389	12	BI346787
32	22	2.0	394	14	CB776687
33	22	2.0	400	9	AA293231
c	34	2.0	402	10	AW136489
35	22	2.0	405	12	BM760479
c	36	2.0	457	10	BE711073
c	37	2.0	464	14	CF308398
c	38	2.0	482	10	AW820357
c	39	2.0	483	10	AW994056
c	40	2.0	491	14	CB725454
41	22	2.0	498	14	W73168
42	22	2.0	512	12	BM830263
c	43	2.0	515	14	CB716451
44	22	2.0	517	9	AA293027
45	22	2.0	523	12	BM841293

#### ALIGNMENTS

```
RESULT 1
BX110362
LOCUS
DEFINITION
  BX110362 NCI CGAP CLL1 Homo sapiens cDNA clone IMAGE5998065203 ;
  IMAGE:2109677, mRNA sequence.
ACCESSION
  BX110362
VERSION
  EST.
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 472)
  Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
  Radlof, U., Schneider, D. and Korn, B.
  Human Unigeneset - RZPD3
  Unpublished (2003)
  Contact: Ina Rolfs
  RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
  Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
  RZPD: IMAGE5998065203
  RZPDLIB: I.M.A.G.E. cDNA Clone Collection;
  Human Unigeneset - RZPD3 (RZPDLIB No.972)
  http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
  RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
  Heubnerweg 6, D-14059 Berlin, Germany
```





```

REFERENCE 1 (bases 1 to 1186)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Günaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 71 Row: m Column: 9
This clone has the following problem: retained intron.
FEATURES
source
1..1186
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4825490"
/tissue_type="Testis"
/clone_lib="NIH MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"
ORIGIN
Query Match 22.9%; Score 250; DB 11; Length 1186;
Best Local Similarity 100.0%; Pred. No. 1.8e-61;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 578 ACCATTCCGGATCTGCTCCAGTGCCTCAACCTCTCCATGCTCCCATGCCACCTGCC 637
Db 802 ACCATTCCGGATCTGCTCCAGTGCCTCAACCTCTCCATGCTCCCATGCCACCTGCC 861
QY 638 ATGTGTGTATCCCGGAGAATACAGAGCAACATGTTGTGTCAGGGGGGCTCCCGGGC 697
Db 862 ATGTGTGTATCCCGGAGAATACAGAGCAACATGTTGTGTCAGGGGGGCTCCCGGGC 921
QY 698 AGGATGCTGCAGGGTGTCTGGGGCCCCCTGTTGTGTGGGGAGTCTTCAAGTTC 757
Db 922 AGGATGCTGCCAGGGTGTCTGGGGCCCCCTGTTGTGTGGGGAGTCTTCAAGTTC 981
QY 758 TGGTGTCTGGGGTCTGTGGGGCCCTGTGACAAAGATGGCATCCCTGGAGTCTACACCT 817
Db 982 TGGTGTCTGGGGTCTGTGGGGCCCTGTGACAAAGATGGCATCCCTGGAGTCTACACCT 1041
QY 818 ATATTGCAA 827
Db 1042 ATATTGCAA 1051
RESULT 4
BG230967
LOCUS BG230967 264 bp mRNA linear EST 09-FEB-2001
DEFINITION nag78a04.x1 NCI CGAP HNS Homo sapiens cDNA clone IMAGE:4227367
similar to TR:Q3UKR1_Q9UKR1 XALLIKREIN-LIKE PROTEIN 5-RELATED
PROTEIN 1. ; mRNA sequence.
ACCESSION BG230967
VERSION BG230967.1 GI:12726030
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 264)
AUTHORS NCI/NIH-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Dental Research,
Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Edward Shillitoe Ph.D., Silvio Gutkind Ph.D.,
Chidchanok Leethanakul D.D.S., Michael Emmert-Buck M.D. Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 257.
FEATURES
source
1..264
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4227367"
/tissue_type="well-differentiated invasive carcinoma,
floor of mouth"
/lab_host="DH10B"
/clone_lib="NCI CGAP HNS"
/note="Vector: PAMPI0; cDNA made by oligo-dT priming.
Non-directionally cloned into the UDG sites of PAMPI0.
Size-selected on agarose gel, average insert size 500 bp.
Primary library; non-amplified. cDNA Library
Preparation: David B. Krizman, Ph.D (NCI). Reference:
Krizman et al. (1996) Cancer Research 56:5380-5383."
ORIGIN
Query Match 19.8%; Score 216; DB 12; Length 264;
Best Local Similarity 100.0%; Pred. No. 2.2e-51;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 857 GGAACTCACTGCTGTTTCTCCACCTCCACCCCACTTAACCTGGGTACCCCTCTG 916
Db 43 GGAACTCACTGCTGTTTCTCCACCTCCACCCCACTTAACCTGGGTACCCCTCTG 102
QY 917 GCCTCAGACGACCAATATCTCTCCATCACTTCCCTAGCTCCTTGTGGCTGG 976
Db 103 GCCTCAGACGACCAATATCTCTCCATCACTTCCCTAGCTCCTTGTGGCTGG 162
QY 977 GAACTTCTTGGAACTTTAACTCTCTGCGAGCCCTTTAAGACCCAGCGGGGTGAGAGA 1036
Db 163 GAACTTCTTGGAACTTTAACTCTCTGCGAGCCCTTTAAGACCCAGCGGGGTGAGAGA 222
QY 1037 AGTGTGCAATAGTCTGGAATAAATAAATGAAGGA 1072
Db 223 AGTGTGCAATAGTCTGGAATAAATAAATGAAGGA 258
RESULT 5
AI554072
LOCUS AI554072 383 bp mRNA linear EST 13-APR-1999
DEFINITION te49g02.x1 Soares NFL T_GBC S1 Homo sapiens cDNA clone
IMAGE:2090066 3', similar to TR:Q99920 Q99920 NES1=NORMAL EPITHELIAL
CELL SPECIFIC GENE 1. ; mRNA sequence.
ACCESSION AI554072
VERSION AI554072.1 GI:4486435
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE 1 (bases 1 to 383)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Insert length: 575 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 373.  
Location/Qualifiers  
FEATURES  
source  
1..383  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2090066"  
/lab\_host="DH10B"  
/clone\_lib="Soares NFL T GBC S1"  
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with  
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;  
Equal amounts of plasmid DNA from three normalized  
libraries (fetal lung NBHL19W, testis NHT, and B-cell  
NCI CGAP GCB1) were mixed, and ss circles were made in  
vitro. Following HAP purification, this DNA was used as  
tracer in a subtractive hybridization reaction. The driver  
was PCR-amplified cDNAs from pools of 5,000 clones made  
from the same 3 libraries. The pools consisted of  
I.M.A.G.E. clones 297480-302087, 682632-687239,  
726408-728711, and 729096-731399. Subtraction by Bento  
Soares and M. Fatima Bonaldo. "  
ORIGIN  
Query Match 9.3%; Score 101; DB 9; Length 383;  
Best Local Similarity 100.0%; Pred. No. 5.9e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 217 CAGCCGTGCGAGTGGGGCTGTTGAGGGCCACAGCGTGGTGGGGGGTGCCTTATT 276  
Db |||||  
QY 192 CAGCCGTGCGAGTGGGGCTGTTGAGGGCCACAGCGTGGTGGGGGGTGCCTTATT 251  
Db |||||  
QY 277 GACCACAGTGGGTCTCTACAGCGGTCTACTGACGGGCAG 317  
Db |||||  
252 GACCACAGTGGGTCTCTACAGCGGTCTACTGACGGGCAG 292  
Db |||||  
RESULT 6  
BX103516/c  
LOCUS BX103516 Soares NFL T GBC S1 Homo sapiens cDNA clone  
DEFINITION IMAGE:2090066; IMAGE:2090066, mRNA sequence.  
ACCESSION BX103516  
VERSION BX103516.1 GI:27945757  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 478)  
Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,  
Radelof, U., Schneider, D. and Korn, B.  
Human Unigeneset - RZPD3  
Unpublished (2003)  
Contact: Ina Rofls  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
RZPD; IMAGE:2090066  
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;  
Human Unigeneset - RZPD3 (RZPDLIB No.972)  
[http://www.rzpd.de/CloneCards/cgi-](http://www.rzpd.de/CloneCards/cgi-bin/showlib.pl.cgi?responsefileNo=972)  
bin/showlib.pl.cgi?responsefileNo=972 Contact: Ina Rofls  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Heubnerweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
www.rzpd.de  
This clone is available royalty-free from RZPD;  
contact RZPD ([clone@rzpd.de](mailto:clone@rzpd.de)) for further information. Seq primer:  
ML3r, Primer sequence: TTTTCACAGGAAACGCTATGAC.  
Location/Qualifiers  
FEATURES  
source  
1..478  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2090066"  
/lab\_host="DH10B"  
/clone\_lib="Soares NFL T GBC S1"  
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with  
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;  
Equal amounts of plasmid DNA from three normalized  
libraries (fetal lung NBHL19W, testis NHT, and B-cell  
NCI CGAP GCB1) were mixed, and ss circles were made in  
vitro. Following HAP purification, this DNA was used as  
tracer in a subtractive hybridization reaction. The driver  
was PCR-amplified cDNAs from pools of 5,000 clones made  
from the same 3 libraries. The pools consisted of  
I.M.A.G.E. clones 297480-302087, 682632-687239,  
726408-728711, and 729096-731399. Subtraction by Bento  
Soares and M. Fatima Bonaldo. "  
ORIGIN  
Query Match 9.3%; Score 101; DB 13; Length 478;  
Best Local Similarity 100.0%; Pred. No. 5e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 217 CAGCCGTGCGAGTGGGGCTGTTGAGGGCCACAGCGTGGTGGGGGGTGCCTTATT 276  
Db |||||  
322 CAGCCGTGCGAGTGGGGCTGTTGAGGGCCACAGCGTGGTGGGGGGTGCCTTATT 263  
QY 277 GACCACAGTGGGTCTCTACAGCGGTCTACTGACGGGCAG 317  
Db |||||  
262 GACCACAGTGGGTCTCTACAGCGGTCTACTGACGGGCAG 222  
Db |||||  
RESULT 7  
BG721150  
LOCUS BG721150 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4825490 5',  
DEFINITION mRNA sequence.  
ACCESSION BG721150  
VERSION BG721150  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 550)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10738 row: 1 column: 03  
High quality sequence stop: 549.  
Location/Qualifiers  
FEATURES  
source  
1..550  
/organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4825490"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
5'-TTTTTTTTTTTTTTNN-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIMH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN
Query Match 4.8%; Score 52; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 GGAAGTGAACCCACCATGGGCTCAGCATCTTTTTCCTGTCGTGTTCTTGG 158
|||||
Db 436 GGAAGTGAACCCACCATGGGCTCAGCATCTTTTTCCTGTCGTGTTCTTGG 487

RESULT 8
BQ129094/c
LOCUS
DEFINITION
BQ129094 577 bp mRNA linear EST 19-APR-2002
i127h02.x1 Melton Amplified Mouse E16 5 Pancreas 3 M16S1 A Mus
musculus cDNA clone IMAGE:5943026 3' similar to TR:Q92IH1 Q92IH1
P6-5 GENE, 3' END ;, mRNA sequence.

ACCESSION
BQ129094
VERSION
BQ129094.1 GI:20203005
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 577)
REFERENCE
AUTHORS
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Scearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarisvili, R.,
Williams, T., Jackson, Y. and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center This clone is
available royalty-free through LLNL; please contact the IMAGE
consortium (info@image.llnl.gov) for further information
Seq primer: -40UP from Gibco
High quality sequence stop: 235.
FEATURES
Location/Qualifiers
source
1..577
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5943026"
/sex="Both"
/tissue_type="pancreas"
/dev_stage="Embryonic day E16.5"
/lab_host="DH10B"
/clone_lib="Melton Amplified Mouse E16 5 Pancreas 3 M16S1

A"
/note="Vector: pSPORT1; Site_1: Not 1; Site 2: Sal 1;
Library constructed using SuperScript plasmid library kit
(Life Technologies). cDNA made by oligo-dT priming.
Size-selected by column fractionation; average insert size
0.97 kb. Amplified once on solid support. cDNA Library
Preparation: Guolin Chen."

ORIGIN
Query Match 2.4%; Score 26; DB 13; Length 577;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 706 TGCACGGGTGATTCTTGGGGCCCCCT 731
|||||
Db 463 TGCACGGGTGATTCTTGGGGCCCCCT 438

RESULT 9
BQ270082
LOCUS
DEFINITION
BQ270082 600 bp mRNA linear EST 15-JUL-2003
ik3302.y1 HR85 islet Homo sapiens cDNA clone IMAGE:5782850 5'
similar to SW:TRY2_HUMAN P07478 TRYPSINOGEN II PRECURSOR ;, mRNA
sequence.

ACCESSION
BQ270082
VERSION
BQ270082.1 GI:20495148
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 600)
REFERENCE
AUTHORS
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Scearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarisvili, R.,
Williams, T., Jackson, Y. and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 422.
FEATURES
Location/Qualifiers
source
1..600
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5782850"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/notes="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

ORIGIN

```

```

Query Match      2.3%; Score 25; DB 13; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 704 CTGCGAGGGTGATTCTGGGGGCC 728
|||||
566 CTGCGAGGGTGATTCTGGGGGCC 590
|||||

RESULT 10
LOCUS BQ269009
DEFINITION ik19g10.y1 HR85 islet Homo sapiens cDNA clone IMAGE:5781690 5',
similar to SW:TRY2_HUMAN P07478 TRYPSINOGEN II PRECURSOR ; mRNA
sequence.
ACCESSION BQ269009
VERSION BQ269009.1 GI:20494075
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 636)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Other ESTs: ik19g10.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 478.
FEATURES
Location/Qualifiers
1..636
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5781690"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

ORIGIN
Query Match      2.3%; Score 25; DB 13; Length 636;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 704 CTGCGAGGGTGATTCTGGGGGCC 728
|||||

```

```

Db 590 CTGCGAGGGTGATTCTGGGGGCC 614

RESULT 11
LOCUS BI794700/c
DEFINITION i663e04.x2 Melton Normalized Mixed Mouse Pancreas 1 N1-MMS1 Mus
musculus cDNA clone IMAGE:5658342 3' similar to SW:EUL_PIG P00772
ELASTASE 1 PRECURSOR ; mRNA sequence.
BI794700
BI794700.1 GI:15822425
EST.
Mus musculus (house mouse)
Mus musculus
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 283)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Other ESTs: ic63e04.y2
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)
Trace considered overall poor quality
MGI:194468 This sequence now available from the IMAGE consortium,
for clone orders contact: info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
FEATURES
Location/Qualifiers
1..283
/organism="Mus musculus"
/mol_type="mRNA"
/strain="ICR"
/db_xref="taxon:10090"
/clone="IMAGE:5658342"
/sex="Both for embryonic & newborn, male for adult and
adult islet"
/dev_stages="Embryonic day 10.5, E12.5, E16.5, newborn,
adult, mixed"
/lab_host="DH10B"
/clone_lib="Melton Normalized Mixed Mouse Pancreas 1
N1-MMS1"
/note="Vector: pSPORT1; Site 1: Not I; Site 2: Sal I; Five
libraries representing E10.5/12.5 pancreatic bud, E16.5
pancreas, newborn pancreas, adult pancreas, and adult
islets of Langerhans were separately constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
was made by oligo-dT priming and size-selected by column
fractionation. Libraries were amplified once on solid
support and plasmid DNA from each library was prepared
and mixed in equal amounts. The mixed library DNA was
normalized by method #4 from Bonaldo, Lennon, and Soares
1996 Genome Research 6:791-806; 0.5 microgram
5 micrograms PCR product representing mixed library
inserts and hybridized to an Ecot of 6. Single-stranded
(unhybridized) plasmids were isolated by hydroxyapatite
chromatography and used to make this library."

```

## ORIGIN

Query Match 2.2%; Score 24; DB 12; Length 283;  
 Best Local Similarity 100.0%; Pred. No. 4.5e+03;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 708 CCAGGCTGATCTCGGGGCCCCCT 731  
 |||||  
 Db 167 CCAGGCTGATCTCGGGGCCCCCT 144

## RESULT 12

AI322408

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AI322408 449 bp mRNA linear EST 23-DEC-1998  
 IMAGE:464085 5' similar to SW:TRY3\_SALSA P35033 TRYPSINOGEN III  
 PRECURSOR ;, mRNA sequence.

AI322408 1 GI:4056837  
 EST.  
 Mus musculus (house mouse)

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 449)  
 Maria, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
 Waterston, R.  
 The WashU-HMI Mouse EST Project  
 Unpublished (1996)  
 Contact: Marra M/Mouse EST Project  
 WashU-HMI Mouse EST Project  
 Washington University School of MedicineP  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.edu  
 This clone is available royalty-free through LLNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MG1:277901  
 This read is a RESEQUENCE of a previously sequenced mouse clone  
 correct orientation)  
 Putative full length read  
 vector to vector length is 607  
 Seq primer: -40Rp from Gibco  
 High quality sequence stop: 408.

Location/Qualifiers  
 1. .449  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:464085"  
 /dev\_stage="19.5 dpc total fetus"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares mouse p3NM19.5"  
 /note="Vector: p773D (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was primed with a Not I - oligo(dT) primer [5'  
 TGTTACCAATCTGAAGTGGCGGCATTTTTTTTTTTTTTTT 3']  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified p773 vector  
 (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by Bento  
 Soares and M. Patricia Bonaldo. RNA was kindly provided by  
 Dr. Minoru Ko (Wayne State University)."

ORIGIN

Query Match

Best Local Similarity

Matches

2.1%; Score 23; DB 9; Length 449;

100.0%; Pred. No. 6.3e+03;

24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1069 AGGAGGGCAAAAAA 1091

Db 24 AGGAGGGCAAAAAA 2

## Matches

23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 594 GTCCAGTGCTCAACCTCTCCA 616  
 |||||  
 Db 147 GTCCAGTGCTCAACCTCTCCA 169

## RESULT 13

CA313295/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CA313295 451 bp mRNA linear EST 04-NOV-2002  
 UI-CF-FNO-aeY-o-05-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone  
 UI-CF-FNO-aeY-o-05-0-UI 3', mRNA sequence.

CA313295  
 EST.  
 Homo sapiens (human)

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 451)  
 Bonaldo, M.F., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)

97044477  
 8889548  
 Contact: McCray, PB  
 McCray Lab  
 University of Iowa  
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
 Tel: 319 356 4866  
 Fax: 319 356 7171  
 Email: paul-mccray@uiowa.edu  
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com) or from Open Biosystems  
 (www.openbiosystems.com).  
 The following repetitive elements were found in this cDNA  
 sequence: 154-233, >(GAA)n#Simple\_repeat  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

Location/Qualifiers  
 1. .451  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-CF-FNO-aeY-o-05-0-UI"  
 /tissue\_type="Human Lung Epithelial cells"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-CF-FNO"  
 /note="Organ: Lung; Vector: p773-Pac (Pharmacia) with a  
 modified polylinker; Site 1: Eco RI; Site 2: Not I;  
 UI-CF-FNO is a subtracted cDNA library derived from two  
 normalized Human lung epithelial cell libraries (EN1 and  
 DUI) the library was subtracted according to according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. For additional information, contact:  
 bento-soares@uiowa.edu  
 TAG TISSUE=Lung Epithelial Cells Tissue nos 359-368  
 TAG\_LIB=UI-CF-FNO  
 TAG\_SEQ=GGCTGTAGGC"

ORIGIN

Query Match

Best Local Similarity

Matches

2.1%; Score 23; DB 14; Length 451;

100.0%; Pred. No. 6.2e+03;

23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1069 AGGAGGGCAAAAAA 1091

Db 24 AGGAGGGCAAAAAA 2

RESULT 14  
AA028356 506 bp mRNA linear EST 11-SEP-1996  
LOCUS m120dli.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone  
DEFINITION IMAGE:464085 5' similar to SW:TRIP\_BOVIN P00760 TRYPSINOGEN ; mRNA sequence.

AA028356  
AA028356.1 GI:1494434

ACCESSION  
VERSION  
KEYWORDS  
SOURCE

Mus musculus (house mouse)

ORGANISM

REFERENCE  
AUTHORS  
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

TITLE  
JOURNAL  
COMMENT

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL ; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

MGI:277901

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 486.

FEATURES

source

1..506

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="IMAGE:464085"

/dev\_stage="19.5 dpc total fetus"

/lab\_host="PH10B (ampicillin resistant)"

/clone\_lib="Soares mouse p3NMF19.5"

/note="Vector: p7T3D (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5',

TGTTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTT 3'],

double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified p7T3 vector

(Pharmacia). Library went through one round of

normalization to a Cot = 5. Library constructed by Bento

Soares and M.Fatima Bonaldo. RNA was kindly provided by

Dr. Minoru Ko (Wayne State University)."

ORIGIN

Query Match

Best Local Similarity

Matches

23; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

2.1%; Score 23; DB 9; Length 506;

100.0%; Pred. No. 5.7e+03;

0; Mismatches

0; Indels

0; Gaps

0;

594 GCTCCAGTGCCTCAACCTCTCCA 616

|||||

146 GCTCCAGTGCCTCAACCTCTCCA 168

|||||

520 bp mRNA linear EST 25-JUN-2003

RESULT 15

EX514743

LOCUS

DEFINITION

EX514743 Soares mouse p3NMF19.5 Mus musculus cDNA clone

IMAGE:464085 ; IMAGE:464085, mRNA sequence.

ACCESSION

EX514743

VERSION

EX514743.1 GI:32244696

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Mus musculus

REFERENCE

AUTHORS

Heil,O., Ebert,L., Neubert,P., Peters,M., Radelof,U., Schneider,D.

and Korn,B.

TITLE

Mus musculus (house mouse)

COMMENT

Unpublished (2003)

Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

RZPD; IMAGp998N221098

RZFDLIB; I.M.A.G.E. cDNA Clone Collection;

Mouse Unigeneset - RZPD2 (RZFDLIB No.981)

http://www.rzpd.de/CloneCards/cgi-

bin/showlib.pl.cgi/response?libNo=981 Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Heubnerweg 5, D-14059 Berlin, Germany

Tel: +49 30 32639 101

Fax: +49 30 32639 111

www.rzpd.de

This clone is available royalty-free from RZPD;

contact RZPD (clone@rzpd.de) for further information. Seq primer:

T7, Primer sequence: TAATACGACTCACTATAGG.

FEATURES

source

1..520

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="IMAGp998N221098 ; IMAGE:464085"

/dev\_stage="19.5 dpc total fetus"

/lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares mouse p3NMF19.5"

/note="Vector: p7T3D (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5',

TGTTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTT 3'],

double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified p7T3 vector

(Pharmacia). Library went through one round of

normalization to a Cot = 5. Library constructed by Bento

Soares and M.Fatima Bonaldo. RNA was kindly provided by

Dr. Minoru Ko (Wayne State University)."

ORIGIN

Query Match

Best Local Similarity

Matches

23; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

2.1%; Score 23; DB 13; Length 520;

100.0%; Pred. No. 5.6e+03;

0; Mismatches

0; Indels

0; Gaps

0;

594 GCTCCAGTGCCTCAACCTCTCCA 616

|||||

148 GCTCCAGTGCCTCAACCTCTCCA 170

|||||

Search completed: June 20, 2004, 09:46:47

Job time : 3358 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 20:26:06 ; Search time 60 Seconds  
(without alignments)  
1167.863 Million cell updates/sec

Title: US-10-015-385A-194

Perfect score: 1374  
Sequence: 1 MGLSIFLLCVLGLSQATP.....GVYTYCKYVDVIRIMVNN 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq 29Jan04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1374	100.0	248	3 AAB21304	AAB21304 Human KIK
2	1374	100.0	248	3 AAB24428	AAB24428 Human PRO
3	1374	100.0	248	3 AAB24032	AAB24032 Human PRO
4	1374	100.0	248	3 AAY99393	AAY99393 Human PRO
5	1374	100.0	248	4 AAM23994	AAM23994 Human EST
6	1374	100.0	248	4 AAB66142	AAB66142 Protein o
7	1374	100.0	248	6 ABO33635	ABO33635 Novel hum
8	1374	100.0	248	7 ABO44488	ABO44488 Human sec
9	1374	100.0	248	7 ABO33512	ABO33512 Novel hum
10	1374	100.0	248	7 ADC18063	ADC18063 Human PRO
11	1374	100.0	248	7 ADD70709	ADD70709 Human sec
12	1374	100.0	248	7 ADD39786	ADD39786 Human sec
13	1374	100.0	248	7 ADD70232	ADD70232 Human sec
14	1374	100.0	248	7 ADD38353	ADD38353 Human sec
15	1374	100.0	248	7 ADD39309	ADD39309 Human sec
16	1374	100.0	248	7 ADD38832	ADD38832 Human sec
17	1374	100.0	248	7 ADD40263	ADD40263 Human sec
18	1374	100.0	248	7 ADE50484	ADE50484 Human sec
19	1374	100.0	248	7 ADE20096	ADE20096 Human sec
20	1374	100.0	248	7 AGE50007	AGE50007 Human sec
21	1374	100.0	248	7 ADE21565	ADE21565 Human sec
22	1367	99.5	248	6 AAO29516	AAO29516 Human kal
23	1301	94.7	254	3 AAB21303	AAB21303 Human KIK
24	1301	94.7	254	5 ABG66676	ABG66676 Human nov
25	1019	74.2	184	3 AAB21301	AAB21301 Human KIK

## ALIGNMENTS

## RESULT 1

ID AAB21304 standard; protein; 248 AA.

XX AAB21304;

DT 02-FEB-2001 (first entry)

XX Human KIK-L5 protein #4.

Human; KIK-L1; KIK-L2; KIK-L3; KIK-L4; KIK-L5; KIK-L6;

KW kallikrein-like protein; serine protease; cytostatic; cancer;

KW prostrate cancer.

XX Homo sapiens.

XX WO200053776-A2.

XX 14-SEP-2000.

XX 09-MAR-2000; 2000WO-CA000258.

XX 11-MAR-1999; 99US-0124260P.

XX 01-APR-1999; 99US-0127386P.

XX 21-JUL-1999; 99US-0144919P.

XX (MOUN ) MOUNT SINAI HOSPITAL.

XX Yousef GM, Diamandis EP;

XX WPI; 2000-587440/55.

XX N-PSDB; AAA95944.

XX New kallikrein-like (KIK-L) proteins for diagnosing and treating KIK-L

XX protein mediated disorders, especially cancer.

XX Claim 12; Page 172; 184pp; English.

XX The present sequence is one of four alternatively spliced kallikrein-like

XX proteins encoded by the human KIK-L4 gene. Kallikreins and kallikrein-

XX like proteins are a subgroup of the serine protease enzyme family. They

XX catalyse the selective cleavage of specific polypeptide precursors to

XX release peptides with potent biological activity. Nucleic acids encoding

XX kallikrein-like proteins KIK-L1, KIK-L2, KIK-L3, KIK-L4, KIK-L5 and KIK-

XX L6 have been isolated. The proteins are useful in the treatment,

XX monitoring and diagnosis of cancers, especially prostate cancer. They

XX can also be used to identify a substance that can associate with or

XX mediate the biological activity of the proteins. Antibodies can be used

Aay28642 Human ser  
Aay32852 Human ser  
Aay41744 Human PRO  
Aay03220 Amino aci  
Aab21322 Human neu  
Aab44300 Human PRO  
Aay51131 Human neu  
Aau12369 Human PRO  
Aab53087 Human ang  
Abg23373 Novel hum  
Abb84852 Human PRO  
Aau81959 Human PRO  
Abb95458 Human ang  
Abo17813 Novel hum  
Abo25246 Novel hum  
Abo81067 Human PRO  
Abo72252 Novel hum  
Ada05702 Human NOV  
Abo66767 Human PRO  
Abo84932 Human sec



CC to treat conditions mediated by the kallikrein-like proteins

```
XX
SQ Sequence 248 AA;
Query Match 100.0%; Score 1374; DB 3; Length 248;
Best Local Similarity 100.0%; Pred. No. 3.8e-98;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLSIFLLLCVLGLSQAATPKIFNGTEGRNSQPMQVGLFEGTSLRCGGVLIDHRWLTA 60
DB 1 MGLSIFLLLCVLGLSQAATPKIFNGTEGRNSQPMQVGLFEGTSLRCGGVLIDHRWLTA 60
QY 61 AHCSGRYVWRLGHEHLSQLDWTQIRHSGFSVTHPGVLGASTSHEHDLRLRLPLVRV 120
DB 61 AHCSGRYVWRLGHEHLSQLDWTQIRHSGFSVTHPGVLGASTSHEHDLRLRLPLVRV 120
QY 121 TSSVQPLPLPNDCATAGTECHVSGWGIINHPRNPPDLLQCLNLSIVSHATCHGVYPGRI 180
DB 121 TSSVQPLPLPNDCATAGTECHVSGWGIINHPRNPPDLLQCLNLSIVSHATCHGVYPGRI 180
QY 181 TSNMVCAGVPGQDQACQDGGPLVCGVGLQGLVSWGSGVPGCGQDGPVYTIKYYVDW 240
DB 181 TSNMVCAGVPGQDQACQDGGPLVCGVGLQGLVSWGSGVPGCGQDGPVYTIKYYVDW 240
QY 241 IRMIMRNN 248
DB 241 IRMIMRNN 248
RESULT 2
AAB24428
ID AAB24428 standard; protein; 248 AA.
AC AAB24428;
DT 07-NOV-2000 (first entry)
DE Human PRO1303 protein sequence SEQ ID NO:203.
XX
KW Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;
KW diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;
KW angiogenic; proliferative; cardiast; cardiovascular; antiatherosclerotic;
KW cytostatic; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
XX WO200032221-A2.
XX
PD 08-JUN-2000.
XX
PF 30-NOV-1999; 99WO-US028313.
XX
XX 01-DEC-1998; 98WO-US025108.
PR 16-DEC-1998; 98US-0112850P.
PR 12-JAN-1999; 99US-0115554P.
PR 08-MAR-1999; 99WO-US005028.
PR 12-MAR-1999; 99US-0123957P.
PR 28-APR-1999; 99US-0131445P.
PR 14-MAY-1999; 99US-0134287P.
PR 02-JUN-1999; 99WO-US012252.
PR 23-JUN-1999; 99US-0141037P.
PR 20-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-OCT-1999; 99US-0162506P.
XX
XX (GETH ) GENENTECH INC.
XX
XX
```

```
PI Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ;
PI Goddard A, Godowski FJ, Gurney AL, Klein RD, Kuo SS, Paoni NF;
PI Smith V, Watanabe CK, Williams PM, Wood WI;
XX
DR WPI: 2000-412154/35.
DR N-PSDB; AAB77671.
XX
PT Nucleic acids encoding PRO polypeptides useful for preventing, diagnosing
PT and treating diagnosing a cardiovascular, endothelial or angiogenic
PT disorders in mammals.
XX
PS Claim 72; Fig 82; 315pp; English.
XX
CC The present invention describes nucleic acids encoding PRO polypeptides
CC useful for preventing, diagnosing and treating diagnosing a
CC cardiovascular, endothelial or angiogenic disorder in mammals by
CC modulating cell proliferation, angiogenesis and cardiovascularisation,
CC and for identifying agonists and antagonists of these processes. The
CC nucleic acids and the proteins they encode may be used in the prevention,
CC treatment and diagnosis of diseases associated with inappropriate PRO
CC expression such as cardiovascular, endothelial or angiogenic disorders in
CC mammals (e.g. atherosclerosis, cancers and cardiac hypertrophy). For
CC example, the nucleic acids (NCs) and vectors containing them and the PRO
CC polypeptide may be used to treat disorders associated with decreased PRO
CC expression. AAB77510 to AAB77721 and AAB24388 to AAB24435 represent
CC nucleotide and protein sequences used in the exemplification of the
CC present invention
XX
SQ Sequence 248 AA;
Query Match 100.0%; Score 1374; DB 3; Length 248;
Best Local Similarity 100.0%; Pred. No. 3.8e-98;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLSIFLLLCVLGLSQAATPKIFNGTEGRNSQPMQVGLFEGTSLRCGGVLIDHRWLTA 60
DB 1 MGLSIFLLLCVLGLSQAATPKIFNGTEGRNSQPMQVGLFEGTSLRCGGVLIDHRWLTA 60
QY 61 AHCSGRYVWRLGHEHLSQLDWTQIRHSGFSVTHPGVLGASTSHEHDLRLRLPLVRV 120
DB 61 AHCSGRYVWRLGHEHLSQLDWTQIRHSGFSVTHPGVLGASTSHEHDLRLRLPLVRV 120
QY 121 TSSVQPLPLPNDCATAGTECHVSGWGIINHPRNPPDLLQCLNLSIVSHATCHGVYPGRI 180
DB 121 TSSVQPLPLPNDCATAGTECHVSGWGIINHPRNPPDLLQCLNLSIVSHATCHGVYPGRI 180
QY 181 TSNMVCAGVPGQDQACQDGGPLVCGVGLQGLVSWGSGVPGCGQDGPVYTIKYYVDW 240
DB 181 TSNMVCAGVPGQDQACQDGGPLVCGVGLQGLVSWGSGVPGCGQDGPVYTIKYYVDW 240
QY 241 IRMIMRNN 248
DB 241 IRMIMRNN 248
RESULT 3
AAB24032
ID AAB24032 standard; protein; 248 AA.
XX
AC AAB24032;
XX
DT 25-JAN-2001 (first entry)
XX
DE Human PRO1303 protein sequence SEQ ID NO:33.
XX
KW Human; tumour; diagnosis; neoplastic disease; proliferation; cancer;
KW identification; tumourigenesis; anticancer; detection.
XX
OS Homo sapiens.
XX
XX WO200053750-A1.
XX
XX 14-SEP-2000.
XX
PD
```

XX PF 02-DEC-1999; 99WO-US028551.  
XX PR 08-MAR-1999; 99WO-US005028.  
XX PR 01-SEP-1999; 99WO-US020111.  
XX PR 29-OCT-1999; 99US-0162506P.  
XX PR 30-NOV-1999; 99WO-US028131.  
XX PR 01-DEC-1999; 99WO-US028634.  
XX PA (GETH ) GENENTECH INC.  
XX PI Botstein D, Goddard A, Gurney AL, Roy MA, Watanabe CK, Wood WI;  
XX DR WPI; 2000-594320/56.  
XX DR N-PSDB; AAC58114.  
XX PT Antibodies specific for PRO polypeptides, used to diagnose and inhibit  
XX PT the growth of tumors in mammals, and to identify inhibitors of PRO  
XX PT polypeptide activity or expression.  
XX PS Claim 61; Fig 24; 226pp; English.  
XX CC The present invention describes an antibody that binds to a human protein  
XX CC (I) selected from: PRO381; PRO1269; PRO1410; PRO1755; PRO1780; PRO2434;  
XX CC PRO1927; PRO3567; PRO1295; PRO1303; PRO4344; PRO4394; PRO4397;  
XX CC PRO4407; PRO1555; PRO1096; PRO2038; and PRO2262. (I) has anticancer  
XX CC activity and can be used to diagnose tumours in mammals, by detecting  
XX CC complex formation when the antibody is contacted with test cells.  
XX CC Increased expression of genes encoding (I) can also be detected to  
XX CC diagnose tumours. Agents which inhibit the activity of (I), especially  
XX CC the antibodies, or an antisense oligonucleotide which hybridises to genes  
XX CC encoding (I), can be used to inhibit tumour growth, preferably by  
XX CC inducing cell death. Methods from the present invention can be used to  
XX CC identify compounds which inhibit the biological activity of (I). AAC58019  
XX CC to AAC58102 represent PCR primers and hybridisation probes used in  
XX CC examples from the present invention for human PRO sequences. AAC58103 to  
XX CC AAC58122 and AAB24021 to AAB24040 represent human PRO polynucleotide and  
XX CC protein sequences given in the exemplification of the present invention  
XX SQ Sequence 248 AA;  
  
Query Match 100.0%; Score 1374; DB 3; Length 248;  
Best Local Similarity 100.0%; Pred. No. 3.8e-98;  
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MGLSIFLLICVLGSLQAATPKIFNGTECGRNSQPMQVGLFEGTSLRCGVLIDHRWLTA 60  
DB 1 MGLSIFLLICVLGSLQAATPKIFNGTECGRNSQPMQVGLFEGTSLRCGVLIDHRWLTA 60  
  
QY 61 AHCSGRVWVRLGSHLSQLDWTEQIRHSGFSVTHPGYLGASTSHEDLRLRLPVRV 120  
DB 61 AHCSGRVWVRLGSHLSQLDWTEQIRHSGFSVTHPGYLGASTSHEDLRLRLPVRV 120  
  
QY 121 TSSVQPLPNDCAATAGTECHVSGWGITVHPRNPDPDLQCLNLSIVSHATGCHVPGRI 180  
DB 121 TSSVQPLPNDCAATAGTECHVSGWGITVHPRNPDPDLQCLNLSIVSHATGCHVPGRI 180  
  
QY 181 TSNMVACAGVPGQDACQSGGSLVCGVGLQGLVSGVSGPCQGDGIPGVYTIKYYVDW 240  
DB 181 TSNMVACAGVPGQDACQSGGSLVCGVGLQGLVSGVSGPCQGDGIPGVYTIKYYVDW 240  
  
QY 241 IRMIMRNN 248  
DB 241 IRMIMRNN 248  
  
RESULT 4  
ID AAY99393  
XX AC AAY99393 standard; protein; 248 AA.  
XX AC AAY99393;  
XX DB 08-AUG-2000 (first entry)

XX DE Human PRO1303 (UNQ669) amino acid sequence SEQ ID NO:194.  
XX KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;  
XX KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening.  
XX OS Homo sapiens.  
XX PN WO200012708-A2.  
XX PD 09-MAR-2000.  
XX PF 01-SEP-1999; 99WO-US020111.  
XX PR 01-SEP-1998; 98US-0098716P.  
XX PR 01-SEP-1998; 98US-0098749P.  
XX PR 02-SEP-1998; 98US-0098750P.  
XX PR 02-SEP-1998; 98US-0098803P.  
XX PR 02-SEP-1998; 98US-0098821P.  
XX PR 02-SEP-1998; 98US-0098843P.  
XX PR 02-SEP-1998; 98US-0098536P.  
XX PR 02-SEP-1998; 98US-0098596P.  
XX PR 02-SEP-1998; 98US-0098598P.  
XX PR 02-SEP-1998; 98US-0098602P.  
XX PR 02-SEP-1998; 98US-0098642P.  
XX PR 10-SEP-1998; 98US-0099741P.  
XX PR 10-SEP-1998; 98US-0099754P.  
XX PR 10-SEP-1998; 98US-0099763P.  
XX PR 10-SEP-1998; 98US-0099792P.  
XX PR 10-SEP-1998; 98US-0099808P.  
XX PR 10-SEP-1998; 98US-0098812P.  
XX PR 10-SEP-1998; 98US-0098815P.  
XX PR 10-SEP-1998; 98US-0098816P.  
XX PR 15-SEP-1998; 98US-0100385P.  
XX PR 15-SEP-1998; 98US-0100388P.  
XX PR 15-SEP-1998; 98US-0100390P.  
XX PR 16-SEP-1998; 98US-0100584P.  
XX PR 16-SEP-1998; 98US-0100627P.  
XX PR 16-SEP-1998; 98US-0100661P.  
XX PR 16-SEP-1998; 98US-0100662P.  
XX PR 16-SEP-1998; 98US-0100664P.  
XX PR 17-SEP-1998; 98US-0100683P.  
XX PR 17-SEP-1998; 98US-0100684P.  
XX PR 17-SEP-1998; 98US-0100710P.  
XX PR 17-SEP-1998; 98US-0100711P.  
XX PR 17-SEP-1998; 98US-0100919P.  
XX PR 17-SEP-1998; 98US-0100930P.  
XX PR 18-SEP-1998; 98US-0100848P.  
XX PR 18-SEP-1998; 98US-0100849P.  
XX PR 18-SEP-1998; 98US-0101044P.  
XX PR 18-SEP-1998; 98US-0101068P.  
XX PR 18-SEP-1998; 98US-0101071P.  
XX PR 22-SEP-1998; 98US-0101279P.  
XX PR 23-SEP-1998; 98US-0101471P.  
XX PR 23-SEP-1998; 98US-0101472P.  
XX PR 23-SEP-1998; 98US-0101473P.  
XX PR 23-SEP-1998; 98US-0101474P.  
XX PR 23-SEP-1998; 98US-0101475P.  
XX PR 23-SEP-1998; 98US-0101476P.  
XX PR 23-SEP-1998; 98US-0101477P.  
XX PR 23-SEP-1998; 98US-0101479P.  
XX PR 24-SEP-1998; 98US-0101738P.  
XX PR 24-SEP-1998; 98US-0101741P.  
XX PR 24-SEP-1998; 98US-0101743P.  
XX PR 24-SEP-1998; 98US-0101915P.  
XX PR 24-SEP-1998; 98US-0101916P.  
XX PR 24-SEP-1998; 98US-0102207P.  
XX PR 29-SEP-1998; 98US-0102240P.  
XX PR 29-SEP-1998; 98US-0102307P.  
XX PR 29-SEP-1998; 98US-0102330P.  
XX PR 29-SEP-1998; 98US-0102331P.  
XX PR 30-SEP-1998; 98US-0102484P.  
XX PR 30-SEP-1998; 98US-0102487P.  
XX PR 30-SEP-1998; 98US-0102570P.

```
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103248P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-0103314P.
PR 07-OCT-1998; 98US-0103315P.
PR 07-OCT-1998; 98US-0103328P.
PR 07-OCT-1998; 98US-0103395P.
PR 07-OCT-1998; 98US-0103396P.
PR 07-OCT-1998; 98US-0103401P.
PR 08-OCT-1998; 98US-0103633P.
PR 08-OCT-1998; 98US-0103678P.
PR 08-OCT-1998; 98US-0103679P.
PR 08-OCT-1998; 98US-0103711P.
PR 14-OCT-1998; 98US-0104257P.
PR 20-OCT-1998; 98US-0104987P.
PR 20-OCT-1998; 98US-0105000P.
PR 20-OCT-1998; 98US-0105002P.
PR 21-OCT-1998; 98US-0105104P.
PR 22-OCT-1998; 98US-0105169P.
PR 22-OCT-1998; 98US-0105266P.
PR 26-OCT-1998; 98US-0105693P.
PR 26-OCT-1998; 98US-0105694P.
PR 27-OCT-1998; 98US-0105807P.
PR 27-OCT-1998; 98US-0105881P.
PR 27-OCT-1998; 98US-0105882P.
PR 28-OCT-1998; 98US-0106062P.
PR 28-OCT-1998; 98US-0106023P.
PR 28-OCT-1998; 98US-0106029P.
PR 28-OCT-1998; 98US-0106030P.
PR 28-OCT-1998; 98US-0106032P.
PR 28-OCT-1998; 98US-0106033P.
PR 29-OCT-1998; 98US-0106178P.
PR 29-OCT-1998; 98US-0106248P.
PR 29-OCT-1998; 98US-0106384P.
PR 29-OCT-1998; 98US-0108500P.
PR 30-OCT-1998; 98US-0106464P.
PR 03-NOV-1998; 98US-0106856P.
PR 03-NOV-1998; 98US-0106902P.
PR 03-NOV-1998; 98US-0106905P.
PR 03-NOV-1998; 98US-0106919P.
PR 03-NOV-1998; 98US-0106932P.
PR 03-NOV-1998; 98US-0106934P.
PR 10-NOV-1998; 98US-0107783P.
PR 17-NOV-1998; 98US-0108775P.
PR 17-NOV-1998; 98US-0108779P.
PR 17-NOV-1998; 98US-0108787P.
PR 17-NOV-1998; 98US-0108801P.
PR 17-NOV-1998; 98US-0108802P.
PR 17-NOV-1998; 98US-0108806P.
PR 17-NOV-1998; 98US-0108807P.
PR 17-NOV-1998; 98US-0108867P.
PR 18-NOV-1998; 98US-0108925P.
PR 18-NOV-1998; 98US-0108848P.
PR 18-NOV-1998; 98US-0108849P.
PR 18-NOV-1998; 98US-0108850P.
PR 18-NOV-1998; 98US-0108851P.
PR 18-NOV-1998; 98US-0108852P.
PR 18-NOV-1998; 98US-0108858P.
PR 18-NOV-1998; 98US-0108904P.
PR
PR (GETH ) GENENTECH INC.
XX
XX Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WT;
XX
XX WPI: 2000-237871/20.
XX N-PSDB; AAA37075.
XX
XX New mammalian DNA sequences encoding transmembrane, receptor or secreted
XX PT PRO polypeptides, useful for screening of potential peptide or small
```

```
PT molecule inhibitors of the relevant receptor/ligand interactions.
XX Claim 12; Fig 108; 773pp; English.
XX
XX AAA37022 to AAA37144 encode the new isolated human transmembrane,
XX receptor or secreted PRO polypeptides given in AA99340 to AA99462. The
XX transmembrane and receptor PRO proteins can be used for screening of
XX potential peptide or small molecule inhibitors of the relevant
XX receptor/ligand interactions. The polypeptides and nucleotide sequences
XX encoding then have various industrial applications, including uses as
XX pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent PCR
XX primers and hybridisation probes used in the isolation of the PRO
XX polypeptides from the present invention
XX
XX Sequence 248 AA;
XX
XX Query Match 100.0%; Score 1374; DB 3; Length 248;
XX Best Local Similarity 100.0%; Pred. No. 3.8e-98;
XX Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MGLSIFLLLCVGLSQAATPKIFNGTECGRNSQPMQVGLFEGTSLRCGVLIDHRWVLT 60
XX Db 1 MGLSIFLLLCVGLSQAATPKIFNGTECGRNSQPMQVGLFEGTSLRCGVLIDHRWVLT 60
XX
XX QY 61 AHCSGSRYWVRLGHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEDLRLRLPYRV 120
XX Db 61 AHCSGSRYWVRLGHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEDLRLRLPYRV 120
XX
XX QY 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHPNRPFDLLQCLNLSIVSHATCHGVYPGR 180
XX Db 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHPNRPFDLLQCLNLSIVSHATCHGVYPGR 180
XX
XX QY 181 TSNMVCAGGVPGQDACQGDSDGPIVCGGVLOGLVSWGSGPCQDGIFGVVTYICKYVDW 240
XX Db 181 TSNMVCAGGVPGQDACQGDSDGPIVCGGVLOGLVSWGSGPCQDGIFGVVTYICKYVDW 240
XX
XX QY 241 IRMINRNN 248
XX Db 241 IRMINRNN 248
XX
XX RESULT 5
XX AAM23994
XX ID AAM23994 standard; protein; 248 AA.
XX
XX AC AAM23994;
XX
XX DT 12-OCT-2001 (first entry)
XX
XX DE Human EST encoded protein SEQ ID NO: 1519.
XX
XX KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
XX tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
XX diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;
XX gene therapy; nutrition.
XX
XX OS Homo sapiens.
XX
XX PN WO200154477-A2.
XX
XX PD 02-AUG-2001.
XX
XX PF 25-JAN-2001; 2001WO-US002687.
XX
XX PR 25-JAN-2000; 2000US-00491404.
XX PR 17-JUL-2000; 2000US-00617746.
XX PR 03-AUG-2000; 2000US-00631451.
XX PR 15-SEP-2000; 2000US-00663870.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
XX Cao Y, Drmanac RA, Zhang J, Werhman T;
```

```

XX DR WPI; 2001-476164/51.
XX DR N-PSDB; AAH98653.
XX PT Isolated polypeptide for treatment of diseases, diagnostics, raising
XX FT antibodies and research use.
XX PS Claim 20; Page 1048-1049; 1275pp; English.
XX CC The present invention provides the protein and coding sequences of novel
XX CC proteins from a variety of organisms, including human, dog, cat, horse,
XX CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
XX CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
XX CC from the organism of interest. They can be used in diagnostics,
XX CC forensics, gene mapping, identification of mutations, to assess
XX CC biodiversity and for nutritional purposes. The present sequence is a
XX CC protein of the invention
XX SQ
XX Query Match 100.0%; Score 1374; DB 4; Length 248;
XX Best Local Similarity 100.0%; Pred. No. 3.8e-98;
XX Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLSIFLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSRRCGVLIDHRWLTA 60
DB 1 MGLSIFLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSRRCGVLIDHRWLTA 60
QY 61 AHCSGSRVWRLGHSLSQLDWTQIRHSGFSVTHPGYLGASTSHEHDLRLRLPLVRV 120
DB 61 AHCSGSRVWRLGHSLSQLDWTQIRHSGFSVTHPGYLGASTSHEHDLRLRLPLVRV 120
QY 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHRPNPFDLLQCLNLSIVSHATCHGVYPGRI 180
DB 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHRPNPFDLLQCLNLSIVSHATCHGVYPGRI 180
QY 181 TSNMVCAGVPGQDACQDGGPLVCGVQLGVLWSVSGVPCGQDGPVVTYICKYVDW 240
DB 181 TSNMVCAGVPGQDACQDGGPLVCGVQLGVLWSVSGVPCGQDGPVVTYICKYVDW 240
QY 241 IRMTMRNN 248
DB 241 IRMTMRNN 248
RESULT 6
AAB66142
ID AAB66142 standard; protein; 248 AA.
AC AAB66142;
XX Protein of the invention #54.
XX Secreted; transmembrane; gene therapy.
XX Unidentified.
XX WO200078961-A1.
XX 28-DEC-2000.
XX 18-FEB-2000; 2000WO-US004342.
XX 23-JUN-1999; 99US-0141037P.
XX 20-JUL-1999; 99US-0144758P.
XX 26-JUL-1999; 99US-0145698P.
XX 01-SEP-1999; 99WO-US020111.
XX 29-OCT-1999; 99US-0162506P.
XX 30-NOV-1999; 99WO-US028313.
XX 02-DEC-1999; 99WO-US028551.
XX 16-DEC-1999; 99WO-US030095.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
XX (GETH ) GENENTECH INC.
XX Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S,
XX Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ,
XX Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK,
XX Williams PM, Wood WI;
XX WPI; 2001-071395/08.
XX Secreted and transmembrane proteins and nucleic acids designated PRO,
XX useful as hybridization probes, in chromosome and gene mapping and gene
XX therapy.
XX Claim 1; Fig 108; 787pp; English.
XX The present invention relates to secreted and transmembrane proteins.
XX These proteins and the DNA encoding them may be used as hybridization
XX probes, in chromosome and gene mapping and in the generation of anti-
XX sense RNA and DNA. They may also be used to generate either
XX transgenic animals or knockout animals which are in turn useful for
XX development and screening of therapeutically useful reagents. The nucleic
XX acids may also be used in gene therapy
XX SQ
XX Query Match 100.0%; Score 1374; DB 4; Length 248;
XX Best Local Similarity 100.0%; Pred. No. 3.8e-98;
XX Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLSIFLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSRRCGVLIDHRWLTA 60
DB 1 MGLSIFLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSRRCGVLIDHRWLTA 60
QY 61 AHCSGSRVWRLGHSLSQLDWTQIRHSGFSVTHPGYLGASTSHEHDLRLRLPLVRV 120
DB 61 AHCSGSRVWRLGHSLSQLDWTQIRHSGFSVTHPGYLGASTSHEHDLRLRLPLVRV 120
QY 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHRPNPFDLLQCLNLSIVSHATCHGVYPGRI 180
DB 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHRPNPFDLLQCLNLSIVSHATCHGVYPGRI 180
QY 181 TSNMVCAGVPGQDACQDGGPLVCGVQLGVLWSVSGVPCGQDGPVVTYICKYVDW 240
DB 181 TSNMVCAGVPGQDACQDGGPLVCGVQLGVLWSVSGVPCGQDGPVVTYICKYVDW 240
QY 241 IRMTMRNN 248
DB 241 IRMTMRNN 248
RESULT 7
ABO33635
ID ABO33635 standard; protein; 248 AA.
XX ABO33635;
XX 17-SEP-2003 (first entry)
XX Novel human secreted and transmembrane protein PRO1303.
XX Human; secreted and transmembrane protein; PRO; angiogenesis;
XX endothelial cell proliferation; wound healing; immune response;
XX T-lymphocytes proliferation; neonatal heart hypertrophy; tumour;
XX cardiac insufficiency disorder; calcium flux; inflammation;
XX vascular endothelial growth factor-stimulated proliferation;
XX mammalian kidney mesangial cell proliferation; Berger disease;
XX nephropathy; Schanlein-Henoch purpura; celiac disease; Crohn's disease;
XX dermatitis herpetiformis; diabetes; haemoglobin switch; insulinemia;
XX pancreatic beta-cell precursor cell differentiation; thalassemias;
XX obesity; auditory hair cell regeneration; hearing loss; bone disorder;

```

KW cartilage disorder; sports injury; arthritis.  
XX Homo sapiens.  
OS US2003073130-A1.  
PN 17-APR-2003.  
XX  
XX 11-DEC-2001; 2001US-00015869.  
XX  
XX 01-SEP-1998; 98US-0098716P.  
XX 01-SEP-1998; 98US-0098723P.  
XX 01-SEP-1998; 98US-0098749P.  
XX 01-SEP-1998; 98US-0098750P.  
XX 02-SEP-1998; 98US-0098803P.  
XX 02-SEP-1998; 98US-0098821P.  
XX 02-SEP-1998; 98US-0098843P.  
XX 02-SEP-1998; 98US-0098536P.  
XX 09-SEP-1998; 98US-0099596P.  
XX 09-SEP-1998; 98US-0099598P.  
XX 09-SEP-1998; 98US-0099602P.  
XX 09-SEP-1998; 98US-0099642P.  
XX 10-SEP-1998; 98US-0099741P.  
XX 10-SEP-1998; 98US-0099754P.  
XX 10-SEP-1998; 98US-0099763P.  
XX 10-SEP-1998; 98US-0099752P.  
XX 10-SEP-1998; 98US-0099808P.  
XX 10-SEP-1998; 98US-0099812P.  
XX 10-SEP-1998; 98US-0099815P.  
XX 15-SEP-1998; 98US-0099816P.  
XX 15-SEP-1998; 98US-0100385P.  
XX 15-SEP-1998; 98US-0100388P.  
XX 15-SEP-1998; 98US-0100390P.  
XX 16-SEP-1998; 98US-0100584P.  
XX 16-SEP-1998; 98US-0100627P.  
XX 16-SEP-1998; 98US-0100661P.  
XX 16-SEP-1998; 98US-0100662P.  
XX 17-SEP-1998; 98US-0100683P.  
XX 17-SEP-1998; 98US-0100684P.  
XX 17-SEP-1998; 98US-0100710P.  
XX 17-SEP-1998; 98US-0100711P.  
XX 17-SEP-1998; 98US-0100919P.  
XX 17-SEP-1998; 98US-0100930P.  
XX 18-SEP-1998; 98US-0100848P.  
XX 18-SEP-1998; 98US-0100849P.  
XX 18-SEP-1998; 98US-0101014P.  
XX 18-SEP-1998; 98US-0101071P.  
XX 18-SEP-1998; 98US-0101078P.  
XX 22-SEP-1998; 98US-0101279P.  
XX 23-SEP-1998; 98US-0101471P.  
XX 23-SEP-1998; 98US-0101472P.  
XX 23-SEP-1998; 98US-0101474P.  
XX 23-SEP-1998; 98US-0101475P.  
XX 23-SEP-1998; 98US-0101476P.  
XX 23-SEP-1998; 98US-0101477P.  
XX 23-SEP-1998; 98US-0101479P.  
XX 24-SEP-1998; 98US-0101738P.  
XX 24-SEP-1998; 98US-0101741P.  
XX 24-SEP-1998; 98US-0101745P.  
XX 24-SEP-1998; 98US-0101915P.  
XX 24-SEP-1998; 98US-0101916P.  
XX 29-SEP-1998; 98US-0102070P.  
XX 29-SEP-1998; 98US-0102240P.  
XX 29-SEP-1998; 98US-0102307P.  
XX 29-SEP-1998; 98US-0102330P.  
XX 30-SEP-1998; 98US-0102331P.  
XX 30-SEP-1998; 98US-0102487P.  
XX 30-SEP-1998; 98US-0102570P.  
XX 30-SEP-1998; 98US-0102571P.  
XX 01-OCT-1998; 98US-0102684P.  
XX 01-OCT-1998; 98US-0102687P.  
XX 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-0103314P.  
PR 07-OCT-1998; 98US-0103315P.  
PR 07-OCT-1998; 98US-0103328P.  
PR 07-OCT-1998; 98US-0103395P.  
PR 07-OCT-1998; 98US-0103396P.  
PR 07-OCT-1998; 98US-0103401P.  
PR 08-OCT-1998; 98US-0103633P.  
PR 08-OCT-1998; 98US-0103678P.  
PR 08-OCT-1998; 98US-0103679P.  
PR 08-OCT-1998; 98US-0103711P.  
PR 14-OCT-1998; 98US-0104257P.  
PR 20-OCT-1998; 98US-0104987P.  
PR 20-OCT-1998; 98US-0105000P.  
PR 20-OCT-1998; 98US-0105002P.  
PR 21-OCT-1998; 98US-0105104P.  
PR 22-OCT-1998; 98US-0105169P.  
PR 22-OCT-1998; 98US-0105266P.  
PR 26-OCT-1998; 98US-0105693P.  
PR 26-OCT-1998; 98US-0105694P.  
PR 27-OCT-1998; 98US-0105807P.  
PR 27-OCT-1998; 98US-0105881P.  
PR 27-OCT-1998; 98US-0105882P.  
PR 27-OCT-1998; 98US-0106062P.  
PR 28-OCT-1998; 98US-0106023P.  
PR 28-OCT-1998; 98US-0106029P.  
PR 28-OCT-1998; 98US-0106030P.  
PR 28-OCT-1998; 98US-0106032P.  
PR 28-OCT-1998; 98US-0106033P.  
PR 28-OCT-1998; 98US-0106178P.  
PR 29-OCT-1998; 98US-0106248P.  
PR 29-OCT-1998; 98US-0106384P.  
PR 29-OCT-1998; 98US-0108500P.  
PR 30-OCT-1998; 98US-0106464P.  
PR 03-NOV-1998; 98US-0106856P.  
PR 03-NOV-1998; 98US-0106902P.  
PR 03-NOV-1998; 98US-0106905P.  
PR 03-NOV-1998; 98US-0106919P.  
PR 03-NOV-1998; 98US-0106932P.  
PR 03-NOV-1998; 98US-0106934P.  
PR 10-NOV-1998; 98US-0107783P.  
PR 17-NOV-1998; 98US-0108775P.  
PR 17-NOV-1998; 98US-0108779P.  
PR 17-NOV-1998; 98US-0108787P.  
PR 17-NOV-1998; 98US-0108788P.  
PR 17-NOV-1998; 98US-0108801P.  
PR 17-NOV-1998; 98US-0108802P.  
PR 17-NOV-1998; 98US-0108806P.  
PR 17-NOV-1998; 98US-0108807P.  
PR 17-NOV-1998; 98US-0108867P.  
PR 17-NOV-1998; 98US-0108925P.  
PR 18-NOV-1998; 98US-0108848P.  
PR 18-NOV-1998; 98US-0108849P.  
PR 18-NOV-1998; 98US-0108850P.  
PR 18-NOV-1998; 98US-0108851P.  
PR 18-NOV-1998; 98US-0108852P.  
PR 18-NOV-1998; 98US-0108858P.  
PR 18-NOV-1998; 98US-0108904P.  
PR 30-DEC-1998; 98US-0113296P.  
PR 30-DEC-1998; 98US-0114223P.  
PR 05-JAN-1999; 99WO-US000106.  
PR 16-APR-1999; 99US-0129674P.  
PR 23-JUN-1999; 99US-0141037P.  
PR 26-JUL-1999; 99US-0144758P.  
PR 26-JUL-1999; 99US-0145698P.  
PR 01-SEP-1999; 99WO-US020111.  
PR 15-SEP-1999; 99WO-US021194.  
PR 29-OCT-1999; 99US-0162506P.  
PR 30-NOV-1999; 99WO-US028313.  
PR 02-DEC-1999; 99WO-US028551.  
PR 16-DEC-1999; 99WO-US030095.  
PR 05-JAN-2000; 2000WO-US000219.



PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102484P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-0103311P.  
PR 07-OCT-1998; 98US-0103315P.  
PR 07-OCT-1998; 98US-0103328P.  
PR 07-OCT-1998; 98US-0103395P.  
PR 07-OCT-1998; 98US-0103396P.  
PR 07-OCT-1998; 98US-0103401P.  
PR 08-OCT-1998; 98US-0103633P.  
PR 08-OCT-1998; 98US-0103678P.  
PR 08-OCT-1998; 98US-0103711P.  
PR 14-OCT-1998; 98US-0104257P.  
PR 20-OCT-1998; 98US-0104987P.  
PR 20-OCT-1998; 98US-0105000P.  
PR 20-OCT-1998; 98US-0105002P.  
PR 21-OCT-1998; 98US-0105104P.  
PR 22-OCT-1998; 98US-0105169P.  
PR 22-OCT-1998; 98US-0105266P.  
PR 26-OCT-1998; 98US-0105693P.  
PR 26-OCT-1998; 98US-0105694P.  
PR 27-OCT-1998; 98US-0105807P.  
PR 27-OCT-1998; 98US-0105881P.  
PR 27-OCT-1998; 98US-0105882P.  
PR 27-OCT-1998; 98US-0106062P.  
PR 28-OCT-1998; 98US-0106023P.  
PR 28-OCT-1998; 98US-0106029P.  
PR 28-OCT-1998; 98US-0106030P.  
PR 28-OCT-1998; 98US-0106032P.  
PR 28-OCT-1998; 98US-0106033P.  
PR 29-OCT-1998; 98US-0106178P.  
PR 29-OCT-1998; 98US-0106248P.  
PR 29-OCT-1998; 98US-0106384P.  
PR 30-OCT-1998; 98US-0108500P.  
PR 30-OCT-1998; 98US-0108464P.  
PR 03-NOV-1998; 98US-0108586P.  
PR 03-NOV-1998; 98US-0106902P.  
PR 03-NOV-1998; 98US-0106905P.  
PR 03-NOV-1998; 98US-0106919P.  
PR 03-NOV-1998; 98US-0106932P.  
PR 03-NOV-1998; 98US-0106934P.  
PR 10-NOV-1998; 98US-0107783P.  
PR 17-NOV-1998; 98US-0108775P.  
PR 17-NOV-1998; 98US-0108779P.  
PR 17-NOV-1998; 98US-0108787P.  
PR 17-NOV-1998; 98US-0108788P.  
PR 17-NOV-1998; 98US-0108801P.  
PR 17-NOV-1998; 98US-0108802P.  
PR 17-NOV-1998; 98US-0108806P.  
PR 17-NOV-1998; 98US-0108807P.  
PR 17-NOV-1998; 98US-0108867P.  
PR 17-NOV-1998; 98US-0108925P.  
PR 18-NOV-1998; 98US-0108846P.  
PR 18-NOV-1998; 98US-0108849P.  
PR 18-NOV-1998; 98US-0108850P.  
PR 18-NOV-1998; 98US-0108851P.  
PR 18-NOV-1998; 98US-0108852P.  
PR 18-NOV-1998; 98US-0108858P.  
PR 18-NOV-1998; 98US-0108904P.  
PR 22-DEC-1998; 98US-0113296P.  
PR 30-DEC-1998; 98US-0114223P.  
PR 05-JAN-1999; 99WO-US000106.  
PR 16-APR-1999; 99US-0129674P.  
PR 23-JUN-1999; 99US-0141037P.  
PR 20-JUL-1999; 99US-0144758P.

PR 26-JUL-1999; 99US-0145698P.  
PR 01-SEP-1999; 99WO-US020111.  
PR 15-SEP-1999; 99WO-US021194.  
PR 29-OCT-1999; 99US-0162506P.  
PR 30-NOV-1999; 99WO-US028313.  
PR 02-DEC-1999; 99WO-US028551.  
PR 16-DEC-1999; 99WO-US030095.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 04-SEP-2001; 2001US-00946374.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S,  
PI Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ,  
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK,  
PI Williams PM, Wood WI;  
XX  
DR WPI; 2003-492259/46.  
DR N-PSDB; ACH04480.  
XX  
PT Novel secreted and transmembrane polypeptides and polynucleotides  
PT encoding them useful for treating various cardiac insufficiency  
PT disorders, bone and/or cartilage disorders such as sports injuries and  
PT arthritis.  
Query Match 100.0%; Score 1374; DB 7; Length 248;  
Best Local Similarity 100.0%; Pred. No. 3.8e-98;  
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGLSIFLLCLVGLGSLQAATPKIFNGTECGRNSQPMQVGLFEGTSLRCGGVLIDHRWLTA 60  
DB 1 MGLSIFLLCLVGLGSLQAATPKIFNGTECGRNSQPMQVGLFEGTSLRCGGVLIDHRWLTA 60  
QY 61 AHCSSGRYVWVRLGEHSLSQLDWTQIRHSGFSVTHPGYLGASTSHEHDLRLRLPLVRV 120  
DB 61 AHCSSGRYVWVRLGEHSLSQLDWTQIRHSGFSVTHPGYLGASTSHEHDLRLRLPLVRV 120  
QY 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHPNRPDPLLOCLNLSIVSHATGCHVPGRI 180  
DB 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHPNRPDPLLOCLNLSIVSHATGCHVPGRI 180  
QY 181 TSNMVCAGGVPQDAGCQDSGGPLVCGGLVQGLVSGSVGPGQDGIPOVYVYICKYVDM 240  
DB 181 TSNMVCAGGVPQDAGCQDSGGPLVCGGLVQGLVSGSVGPGQDGIPOVYVYICKYVDM 240  
QY 241 IRMIMENN 248  
DB 241 IRMIMENN 248  
RESULT 9  
ABO33512

ID ABO33512 standard; protein; 248 AA.  
XX  
AC ABO33512;  
XX  
DT 17-SEP-2003 (first entry)  
XX  
DE Novel human secreted and transmembrane protein PRO1303.  
XX  
KW Human; secreted and transmembrane protein; PRO; gene therapy; vaccine;  
KW tissue typing; chromosome identification; vaccine.  
OS Homo sapiens.  
XX  
PN US2003073129-A1.  
XX  
PD 17-APR-2003.  
XX  
XX 04-SEP-2001; 2001US-00946374.  
XX  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 01-SEP-1998; 98US-0098749P.  
PR 01-SEP-1998; 98US-0098750P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099536P.  
PR 09-SEP-1998; 98US-0099596P.  
PR 09-SEP-1998; 98US-0099598P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 09-SEP-1998; 98US-0099642P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099792P.  
PR 10-SEP-1998; 98US-0099808P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 10-SEP-1998; 98US-0099815P.  
PR 10-SEP-1998; 98US-0099818P.  
PR 15-SEP-1998; 98US-0100385P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 15-SEP-1998; 98US-0100390P.  
PR 16-SEP-1998; 98US-0100584P.  
PR 16-SEP-1998; 98US-0100627P.  
PR 16-SEP-1998; 98US-0100661P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100710P.  
PR 17-SEP-1998; 98US-0100711P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100848P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 18-SEP-1998; 98US-0101071P.  
PR 22-SEP-1998; 98US-0101279P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101474P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101476P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 23-SEP-1998; 98US-0101479P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101741P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101915P.  
PR 24-SEP-1998; 98US-0101916P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102307P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102484P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-0103314P.  
PR 07-OCT-1998; 98US-0103315P.  
PR 07-OCT-1998; 98US-0103328P.  
PR 07-OCT-1998; 98US-0103395P.  
PR 07-OCT-1998; 98US-0103396P.  
PR 07-OCT-1998; 98US-0103401P.  
PR 08-OCT-1998; 98US-0103633P.  
PR 08-OCT-1998; 98US-0103678P.  
PR 08-OCT-1998; 98US-0103679P.  
PR 08-OCT-1998; 98US-0103711P.  
PR 14-OCT-1998; 98US-0104257P.  
PR 20-OCT-1998; 98US-0104987P.  
PR 20-OCT-1998; 98US-0105000P.  
PR 20-OCT-1998; 98US-0105082P.  
PR 21-OCT-1998; 98US-0105104P.  
PR 22-OCT-1998; 98US-0105169P.  
PR 22-OCT-1998; 98US-0105266P.  
PR 26-OCT-1998; 98US-0105693P.  
PR 26-OCT-1998; 98US-0105694P.  
PR 27-OCT-1998; 98US-0105807P.  
PR 27-OCT-1998; 98US-0105881P.  
PR 27-OCT-1998; 98US-0105882P.  
PR 28-OCT-1998; 98US-0106023P.  
PR 28-OCT-1998; 98US-0106029P.  
PR 28-OCT-1998; 98US-0106030P.  
PR 28-OCT-1998; 98US-0106032P.  
PR 28-OCT-1998; 98US-0106033P.  
PR 28-OCT-1998; 98US-0106178P.  
PR 29-OCT-1998; 98US-0106248P.  
PR 29-OCT-1998; 98US-0106384P.  
PR 29-OCT-1998; 98US-0108500P.  
PR 30-OCT-1998; 98US-0106484P.  
PR 03-NOV-1998; 98US-0106856P.  
PR 03-NOV-1998; 98US-0106902P.  
PR 03-NOV-1998; 98US-0106905P.  
PR 03-NOV-1998; 98US-0106919P.  
PR 03-NOV-1998; 98US-0106932P.  
PR 03-NOV-1998; 98US-0106934P.  
PR 10-NOV-1998; 98US-0107783P.  
PR 17-NOV-1998; 98US-0108775P.  
PR 17-NOV-1998; 98US-0108779P.  
PR 17-NOV-1998; 98US-0108787P.  
PR 17-NOV-1998; 98US-0108788P.  
PR 17-NOV-1998; 98US-0108801P.  
PR 17-NOV-1998; 98US-0108802P.  
PR 17-NOV-1998; 98US-0108806P.  
PR 17-NOV-1998; 98US-0108807P.  
PR 17-NOV-1998; 98US-0108867P.  
PR 18-NOV-1998; 98US-0108925P.  
PR 18-NOV-1998; 98US-0108848P.  
PR 18-NOV-1998; 98US-0108849P.  
PR 18-NOV-1998; 98US-0108850P.  
PR 18-NOV-1998; 98US-0108851P.  
PR 18-NOV-1998; 98US-0108852P.  
PR 18-NOV-1998; 98US-0108858P.  
PR 18-NOV-1998; 98US-0108904P.  
PR 22-DEC-1998; 98US-00218517.  
PR 22-DEC-1998; 98US-0113296P.  
PR 30-DEC-1998; 98US-0114223P.  
PR 05-JAN-1999; 99WO-US000106.



```
PR 12-APR-1999; 99US-00284291.
PR 16-APR-1999; 99US-0129674P.
PR 23-JUN-1999; 99US-0141037P.
PR 20-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021194.
PR 18-OCT-1999; 99US-00403297.
PR 30-NOV-1999; 99WO-US028313.
PR 02-DEC-1999; 99WO-US028551.
PR 16-DEC-1999; 99WO-US030095.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US000365.
PR 18-FEB-2000; 2000WO-US004342.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006884.
PR 17-MAY-2000; 2000WO-US0113705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 01-JUN-2001; 2001US-00872035.
PR 14-JUN-2001; 2001WO-US017800.
PR 14-JUN-2001; 2001US-00882636.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
XX Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
XX Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
XX Williams PM, Wood W;
XX
XX WPI: 2003-585292/55.
XX N-PSDB; ACD68024.
XX
XX Novel isolated PRO polypeptides e.g. PRO1491 and PRO1571, useful in the
XX preparation of a medicament for treating a condition responsive to PRO
XX polypeptide, and as therapeutic agents e.g. vaccines.
XX
XX Claim 12; Fig 108; 561pp; English.
XX
XX The invention describes an isolated PRO (secreted and transmembrane)
XX polypeptide (I), having at least 80% sequence identity to a sequence
XX
XX Query Match 100.0%; Score 1374; DB 7; Length 248;
XX Best Local Similarity 100.0%; Pred. No. 3.8e-98;
XX Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MGLSIFLLLCVLGSLQAATPKIFNGTECGRNSQPMQVGLFEGTSLRCGGVLIDHRWVLT 60
Db 1 MGLSIFLLLCVLGSLQAATPKIFNGTECGRNSQPMQVGLFEGTSLRCGGVLIDHRWVLT 60
QY 61 AHCSGSRYWRLGHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLPVRV 120
Db 61 AHCSGSRYWRLGHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLPVRV 120
QY 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHPNPFPPDLLQCLNLSIVSHATCHGVPGRI 180
Db 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHPNPFPPDLLQCLNLSIVSHATCHGVPGRI 180
QY 181 TSNMVCAGGVPQDACCQDGGPLVCGGV.LQGLVSWGSGVPCQGDGIPGVYTYICKYVDW 240
Db 181 TSNMVCAGGVPQDACCQDGGPLVCGGV.LQGLVSWGSGVPCQGDGIPGVYTYICKYVDW 240
```

```
Db 181 TSNMVCAGGVPQDACCQDGGPLVCGGV.LQGLVSWGSGVPCQGDGIPGVYTYICKYVDW 240
QY 241 IIRMIMENN 248
Db 241 IIRMIMENN 248

RESULT 10
ADC18063
ID ADC18063 standard; protein; 248 AA.
XX
XX AC ADC18063;
XX
XX DT 18-DEC-2003 (first entry)
XX
XX DE Human PRO polypeptide #54.
XX
XX KW Human; PRO; protein electrophoresis; chromosome mapping; gene mapping;
XX genetic disorder.
XX
XX OS Homo sapiens.
XX
XX PN US2003064925-A1.
XX
XX PD 03-APR-2003.
XX
XX PF 10-DEC-2001; 2001US-00013907.
XX
XX PR 01-SEP-1998; 98US-0098716P.
XX PR 01-SEP-1998; 98US-0098723P.
XX PR 01-SEP-1998; 98US-0098749P.
XX PR 01-SEP-1998; 98US-0098750P.
XX PR 02-SEP-1998; 98US-0098803P.
XX PR 02-SEP-1998; 98US-0098821P.
XX PR 02-SEP-1998; 98US-0098843P.
XX PR 09-SEP-1998; 98US-0099536P.
XX PR 09-SEP-1998; 98US-0099596P.
XX PR 09-SEP-1998; 98US-0099598P.
XX PR 09-SEP-1998; 98US-0099602P.
XX PR 09-SEP-1998; 98US-0099642P.
XX PR 10-SEP-1998; 98US-0099741P.
XX PR 10-SEP-1998; 98US-0099754P.
XX PR 10-SEP-1998; 98US-0099763P.
XX PR 10-SEP-1998; 98US-0099792P.
XX PR 10-SEP-1998; 98US-0099808P.
XX PR 10-SEP-1998; 98US-0099812P.
XX PR 10-SEP-1998; 98US-0099815P.
XX PR 10-SEP-1998; 98US-0099816P.
XX PR 15-SEP-1998; 98US-0100385P.
XX PR 15-SEP-1998; 98US-0100388P.
XX PR 15-SEP-1998; 98US-0100390P.
XX PR 16-SEP-1998; 98US-0100584P.
XX PR 16-SEP-1998; 98US-0100627P.
XX PR 16-SEP-1998; 98US-0100661P.
XX PR 16-SEP-1998; 98US-0100662P.
XX PR 16-SEP-1998; 98US-0100664P.
XX PR 17-SEP-1998; 98US-0100683P.
XX PR 17-SEP-1998; 98US-0100684P.
XX PR 17-SEP-1998; 98US-0100710P.
XX PR 17-SEP-1998; 98US-0100711P.
XX PR 17-SEP-1998; 98US-0100919P.
XX PR 17-SEP-1998; 98US-0100930P.
XX PR 18-SEP-1998; 98US-0100848P.
XX PR 18-SEP-1998; 98US-0100849P.
XX PR 18-SEP-1998; 98US-0101014P.
XX PR 18-SEP-1998; 98US-0101068P.
XX PR 22-SEP-1998; 98US-0101279P.
XX PR 23-SEP-1998; 98US-0101471P.
XX PR 23-SEP-1998; 98US-0101472P.
XX PR 23-SEP-1998; 98US-0101474P.
XX PR 23-SEP-1998; 98US-0101475P.
XX PR 23-SEP-1998; 98US-0101476P.
```

```
PR 23-SEP-1998; 98US-0101477P.
PR 23-SEP-1998; 98US-0101478P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101741P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101915P.
PR 24-SEP-1998; 98US-0101916P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102307P.
PR 29-SEP-1998; 98US-0102330P.
PR 30-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102484P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-0103314P.
PR 07-OCT-1998; 98US-0103315P.
PR 07-OCT-1998; 98US-0103328P.
PR 07-OCT-1998; 98US-0103395P.
PR 07-OCT-1998; 98US-0103396P.
PR 07-OCT-1998; 98US-0103401P.
PR 08-OCT-1998; 98US-0103633P.
PR 08-OCT-1998; 98US-0103678P.
PR 08-OCT-1998; 98US-0103679P.
PR 08-OCT-1998; 98US-0103711P.
PR 14-OCT-1998; 98US-0104257P.
PR 20-OCT-1998; 98US-0104987P.
PR 20-OCT-1998; 98US-0105000P.
PR 20-OCT-1998; 98US-0105002P.
PR 21-OCT-1998; 98US-0105104P.
PR 22-OCT-1998; 98US-0105169P.
PR 22-OCT-1998; 98US-0105266P.
PR 26-OCT-1998; 98US-0105693P.
PR 26-OCT-1998; 98US-0105694P.
PR 27-OCT-1998; 98US-0105807P.
PR 27-OCT-1998; 98US-0105881P.
PR 27-OCT-1998; 98US-0105882P.
PR 27-OCT-1998; 98US-0106062P.
PR 28-OCT-1998; 98US-0106023P.
PR 28-OCT-1998; 98US-0106029P.
PR 28-OCT-1998; 98US-0106030P.
PR 28-OCT-1998; 98US-0106032P.
PR 28-OCT-1998; 98US-0106033P.
PR 29-OCT-1998; 98US-0106178P.
PR 29-OCT-1998; 98US-0106248P.
PR 29-OCT-1998; 98US-0106384P.
PR 29-OCT-1998; 98US-0108500P.
PR 30-OCT-1998; 98US-0106464P.
PR 03-NOV-1998; 98US-0106856P.
PR 03-NOV-1998; 98US-0106902P.
PR 03-NOV-1998; 98US-0106905P.
PR 03-NOV-1998; 98US-0106919P.
PR 03-NOV-1998; 98US-0106932P.
PR 03-NOV-1998; 98US-0106934P.
PR 10-NOV-1998; 98US-0107783P.
PR 17-NOV-1998; 98US-0108775P.
PR 17-NOV-1998; 98US-0108779P.
PR 17-NOV-1998; 98US-0108787P.
PR 17-NOV-1998; 98US-0108788P.
PR 17-NOV-1998; 98US-0108801P.
PR 17-NOV-1998; 98US-0108802P.
PR 17-NOV-1998; 98US-0108806P.
PR 17-NOV-1998; 98US-0108807P.
PR 17-NOV-1998; 98US-0108867P.
PR 17-NOV-1998; 98US-0108925P.
PR 18-NOV-1998; 98US-0108848P.
PR 18-NOV-1998; 98US-0108849P.

PR 18-NOV-1998; 98US-0108850P.
PR 18-NOV-1998; 98US-0108851P.
PR 18-NOV-1998; 98US-0108852P.
PR 18-NOV-1998; 98US-0108858P.
PR 18-NOV-1998; 98US-0108904P.
PR 22-DEC-1998; 98US-0113296P.
PR 30-DEC-1998; 98US-0114223P.
PR 05-JAN-1999; 99WO-US000106.
PR 16-APR-1999; 99US-0129674P.
PR 23-JUN-1999; 99US-0141037P.
PR 20-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021194.
PR 29-OCT-1999; 99US-0162506P.
PR 30-NOV-1999; 99WO-US028313.
PR 02-DEC-1999; 99WO-US028551.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004342.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006884.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006656.
PR 01-JUN-2001; 2001WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 04-SEP-2001; 2001US-00946374.

( GETH ) GENENTECH INC.
XX
XX
XX
PI Baker KP, Botstein D, Desnovers L, Eaton DL, Ferrara N, Fong S;
PI Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
PI Williams PM, Wood WI;
XX
XX WPI; 2003-555602/52.
DR N-PSDB; ADC18062.
XX
XX
XX Novel isolated PRO polypeptides e.g. PRO1491 and PRO1571, useful in the
PT preparation of a medicament for treating a condition responsive to PRO
PT polypeptide, and as therapeutic agents e.g. vaccines.
XX
XX Claim 12; SEQ ID NO 194; 555pp; English.
XX
XX The invention relates to human PRO polypeptides and the polynucleotides
CC encoding them. The sequences are useful in the preparation of a
CC medicament for treating a condition responsive to a PRO polypeptide. The
CC polypeptides are useful in a number of functional biological assays, as
CC molecular weight markers for protein electrophoresis and as therapeutic

Query Match 100.0%; Score 1374; DB 7; Length 248;
Best Local Similarity 100.0%; Pred. No. 3.8e-98;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLSIFLLLCVLGSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWLTA 60
Db 1 MGLSIFLLLCVLGSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWLTA 60
Qy 61 AHCSGRYVVRNLGHSLSQLDWTQIRHSGFSVTHPGVLGASTSHEHDLRLRLPVRV 120
```

```
Db      61  AHCSGSRWYRLGHSLSQDWTQIRHSGFSVTHPGYLGASTSHEHDLRLRLLEVRV 120
Qy      121  TSSVQPLPLPNDACATAGTECHVSGWGITNHPRNPFPDLLQCLNLSIVSHATCHGVYPGRI 180
Db      121  TSSVQPLPLPNDACATAGTECHVSGWGITNHPRNPFPDLLQCLNLSIVSHATCHGVYPGRI 180
Qy      181  TSNMVCAGGVPQDACQDGGPLVCGGVLOGLVSGVSGVPGCGQDGPVGVYTYICKYVDW 240
Db      181  TSNMVCAGGVPQDACQDGGPLVCGGVLOGLVSGVSGVPGCGQDGPVGVYTYICKYVDW 240
Qy      241  IRMIMRNN 248
Db      241  IRMIMRNN 248

RESULT 11
ADD70709
ID  ADD70709 standard; protein; 248 AA.
XX  AC  ADD70709;
XX  DT  15-JAN-2004 (first entry)
XX  DE  Human secreted/transmembrane protein PRO1303.
XX  KW  Human; secreted protein; transmembrane protein; PRO; tumour;
KW  immune response; cardiac insufficiency disorder; calcium flux;
KW  umbilical vein endothelial cell; bone disorder; cartilage disorder;
KW  arthritis; wound healing; diabetes; skeletal muscle cells; obesity;
KW  Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;
KW  dermatitis; herpeticiformis; Crohn's disease; thalassaemia.
XX  OS  Homo sapiens.
XX  SS  US2003099625-A1.
XX  PD  29-MAY-2003.
XX  PF  12-DEC-2001; 2001US-00015386.
XX  PR  01-SEP-1998; 98US-0098716P.
XX  PR  01-SEP-1998; 98US-0098723P.
XX  PR  01-SEP-1998; 98US-0098743P.
XX  PR  01-SEP-1998; 98US-0098750P.
XX  PR  02-SEP-1998; 98US-0098803P.
XX  PR  02-SEP-1998; 98US-0098821P.
XX  PR  02-SEP-1998; 98US-0098843P.
XX  PR  09-SEP-1998; 98US-0099536P.
XX  PR  09-SEP-1998; 98US-0099596P.
XX  PR  09-SEP-1998; 98US-0099602P.
XX  PR  09-SEP-1998; 98US-0099642P.
XX  PR  10-SEP-1998; 98US-0099741P.
XX  PR  10-SEP-1998; 98US-0099754P.
XX  PR  10-SEP-1998; 98US-0099763P.
XX  PR  10-SEP-1998; 98US-0099792P.
XX  PR  10-SEP-1998; 98US-0099808P.
XX  PR  10-SEP-1998; 98US-0099812P.
XX  PR  10-SEP-1998; 98US-0099815P.
XX  PR  15-SEP-1998; 98US-0099816P.
XX  PR  15-SEP-1998; 98US-0100385P.
XX  PR  15-SEP-1998; 98US-0100388P.
XX  PR  15-SEP-1998; 98US-0100390P.
XX  PR  16-SEP-1998; 98US-0100584P.
XX  PR  16-SEP-1998; 98US-0100627P.
XX  PR  16-SEP-1998; 98US-0100661P.
XX  PR  16-SEP-1998; 98US-0100662P.
XX  PR  16-SEP-1998; 98US-0100664P.
XX  PR  17-SEP-1998; 98US-0100683P.
XX  PR  17-SEP-1998; 98US-0100684P.
XX  PR  17-SEP-1998; 98US-0100710P.
XX  PR  17-SEP-1998; 98US-0100711P.
PR  17-SEP-1998; 98US-0100919P.
PR  17-SEP-1998; 98US-0100930P.
PR  18-SEP-1998; 98US-0100848P.
PR  18-SEP-1998; 98US-0101014P.
PR  18-SEP-1998; 98US-0101068P.
PR  18-SEP-1998; 98US-0101071P.
PR  22-SEP-1998; 98US-0101279P.
PR  23-SEP-1998; 98US-0101471P.
PR  23-SEP-1998; 98US-0101472P.
PR  23-SEP-1998; 98US-0101474P.
PR  23-SEP-1998; 98US-0101475P.
PR  23-SEP-1998; 98US-0101476P.
PR  23-SEP-1998; 98US-0101477P.
PR  23-SEP-1998; 98US-0101479P.
PR  24-SEP-1998; 98US-0101738P.
PR  24-SEP-1998; 98US-0101741P.
PR  24-SEP-1998; 98US-0101743P.
PR  24-SEP-1998; 98US-0101915P.
PR  24-SEP-1998; 98US-0101916P.
PR  29-SEP-1998; 98US-0102207P.
PR  29-SEP-1998; 98US-0102240P.
PR  29-SEP-1998; 98US-0102307P.
PR  29-SEP-1998; 98US-0102330P.
PR  29-SEP-1998; 98US-0102331P.
PR  30-SEP-1998; 98US-0102484P.
PR  30-SEP-1998; 98US-0102487P.
PR  30-SEP-1998; 98US-0102570P.
PR  01-OCT-1998; 98US-0102571P.
PR  01-OCT-1998; 98US-0102684P.
PR  01-OCT-1998; 98US-0102687P.
PR  02-OCT-1998; 98US-0102965P.
PR  06-OCT-1998; 98US-0103258P.
PR  06-OCT-1998; 98US-0103449P.
PR  07-OCT-1998; 98US-0103314P.
PR  07-OCT-1998; 98US-0103315P.
PR  07-OCT-1998; 98US-0103328P.
PR  07-OCT-1998; 98US-0103395P.
PR  07-OCT-1998; 98US-0103396P.
PR  07-OCT-1998; 98US-0103401P.
PR  08-OCT-1998; 98US-0103633P.
PR  08-OCT-1998; 98US-0103678P.
PR  08-OCT-1998; 98US-0103679P.
PR  08-OCT-1998; 98US-0103711P.
PR  14-OCT-1998; 98US-0104257P.
PR  20-OCT-1998; 98US-0104987P.
PR  20-OCT-1998; 98US-0105000P.
PR  20-OCT-1998; 98US-0105002P.
PR  21-OCT-1998; 98US-0105104P.
PR  22-OCT-1998; 98US-0105169P.
PR  22-OCT-1998; 98US-0105266P.
PR  26-OCT-1998; 98US-0105693P.
PR  26-OCT-1998; 98US-0105694P.
PR  27-OCT-1998; 98US-0105807P.
PR  27-OCT-1998; 98US-0105881P.
PR  27-OCT-1998; 98US-0105882P.
PR  27-OCT-1998; 98US-0106062P.
PR  28-OCT-1998; 98US-0106023P.
PR  28-OCT-1998; 98US-0106029P.
PR  28-OCT-1998; 98US-0106030P.
PR  28-OCT-1998; 98US-0106032P.
PR  28-OCT-1998; 98US-0106033P.
PR  28-OCT-1998; 98US-0106178P.
PR  29-OCT-1998; 98US-0106248P.
PR  29-OCT-1998; 98US-0106384P.
PR  29-OCT-1998; 98US-0108500P.
PR  30-OCT-1998; 98US-0106464P.
PR  03-NOV-1998; 98US-0106856P.
PR  03-NOV-1998; 98US-0106902P.
PR  03-NOV-1998; 98US-0106905P.
PR  03-NOV-1998; 98US-0106919P.
PR  03-NOV-1998; 98US-0106932P.
PR  03-NOV-1998; 98US-0106934P.
```

PR 10-NOV-1998;	98US-0107783P.	Query Match	100.0%;	Score 1374;	DB 7;	Length 248;
PR 17-NOV-1998;	98US-0108775P.	Best Local Similarity	100.0%;	Pred. No. 3.8e-98;		
PR 17-NOV-1998;	98US-0108775P.	Matches 248;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
PR 17-NOV-1998;	98US-0108787P.					
PR 17-NOV-1998;	98US-0108788P.	Qy 1	MGLSIFLLLCVGLSQAATPKIFNGTECGRNSQFQVGLFEGTSLRCGGVLIDHRWLTA	60		
PR 17-NOV-1998;	98US-0108801P.	Db 1	MGLSIFLLLCVGLSQAATPKIFNGTECGRNSQFQVGLFEGTSLRCGGVLIDHRWLTA	60		
PR 17-NOV-1998;	98US-0108802P.					
PR 17-NOV-1998;	98US-0108806P.					
PR 17-NOV-1998;	98US-0108807P.	Qy 61	AHCSGSRYWVRLGEHSLSQLDWTEQIRHSGFVTHPGYLGASTSHEHDLRLRLPRV	120		
PR 17-NOV-1998;	98US-0108825P.	Db 61	AHCSGSRYWVRLGEHSLSQLDWTEQIRHSGFVTHPGYLGASTSHEHDLRLRLPRV	120		
PR 18-NOV-1998;	98US-0108848P.	Qy 121	TSSVQPLPLPNDCATAGTECHVSGWGITNHPRPFPDLLQCLNLSIVSHATCHGVYGR	180		
PR 18-NOV-1998;	98US-0108849P.	Db 121	TSSVQPLPLPNDCATAGTECHVSGWGITNHPRPFPDLLQCLNLSIVSHATCHGVYGR	180		
PR 18-NOV-1998;	98US-0108850P.					
PR 18-NOV-1998;	98US-0108851P.	Qy 181	TSMNVCAGGVPGQDACQDGGPLVCGGVLCGLVSWGSGVPCGQDGPVYTYICKYVDW	240		
PR 18-NOV-1998;	98US-0108852P.	Db 181	TSMNVCAGGVPGQDACQDGGPLVCGGVLCGLVSWGSGVPCGQDGPVYTYICKYVDW	240		
PR 18-NOV-1998;	98US-0108904P.					
PR 22-DEC-1998;	98US-0113296P.	Qy 241	IRMIMRNN 248			
PR 30-DEC-1998;	98US-0114223P.	Db 241	IRMIMRNN 248			
PR 05-JAN-1999;	99WO-US000106.					
PR 15-SEP-1999;	99WO-US021194.					
PR 29-OCT-1999;	99US-0129674P.					
PR 23-JUN-1999;	99US-0141037P.					
PR 20-JUL-1999;	99US-0144758P.					
PR 26-JUL-1999;	99US-0145698P.					
PR 01-SEP-1999;	99WO-US020111.					
PR 15-SEP-1999;	99WO-US021194.					
PR 30-NOV-1999;	99WO-US028313.					
PR 02-DEC-1999;	99WO-US028351.					
PR 16-DEC-1999;	99WO-US030095.					
PR 05-JAN-2000;	2000WO-US000219.					
PR 06-JAN-2000;	2000WO-US000376.					
PR 11-FEB-2000;	2000WO-US003565.					
PR 18-FEB-2000;	2000WO-US004342.					
PR 24-FEB-2000;	2000WO-US005004.					
PR 02-MAR-2000;	2000WO-US005841.					
PR 15-MAR-2000;	2000WO-US006884.					
PR 17-MAY-2000;	2000WO-US013705.					
PR 22-MAY-2000;	2000WO-US014042.					
PR 30-MAY-2000;	2000WO-US014941.					
PR 02-JUN-2000;	2000WO-US015264.					
PR 23-AUG-2000;	2000WO-US023522.					
PR 24-AUG-2000;	2000WO-US023328.					
PR 08-NOV-2000;	2000WO-US030952.					
PR 10-NOV-2000;	2000WO-US030873.					
PR 01-DEC-2000;	2000WO-US032678.					
PR 28-FEB-2001;	2001WO-US006520.					
XX						
PA	(GETH ) GENENTECH INC.					
XX	Baker KP, Botstein D, Deenoyers L, Eaton DL, Ferrara N, Fong S;					
PI	Gao W, Goddard A, Godowski J, Grimaldi JC, Gurney AL, Hillan KJ;					
PI	Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;					
PI	Williams PM, Wood WI;					
XX						
DR	WPI: 2003-874602/81.					
DR	N-PSDB; ADD70708.					
XX						
PT	Novel isolated PRO polypeptides e.g., PRO1130, PRO1275, PRO1418, PRO1555,					
PT	PRO1787 affect glucose or free fatty acid (FFA) uptake by skeletal muscle					
PT	cells and are useful for treating diabetes or hyper- or hypo-insulinemia.					
XX						
PS	Claim 12; SEQ ID NO 194; 53pp; English.					
XX						
CC	The invention relates to an isolated PRO polypeptide (secreted or					

RESULT 12

ADD33786

ID ADD33786 standard; protein; 248 AA.

XX AC ADD33786;

XX DT 15-JAN-2004 (first entry)

XX DE Human secreted/transmembrane protein PRO1303.

XX KW Human; secreted protein; transmembrane protein; PRO; tumour;  
 KW immune response; cardiac insufficiency disorder; calcium flux;  
 KW umbilical vein endothelial cell; bone disorder; cartilage disorder;  
 KW arthritis; wound healing; diabetes; skeletal muscle cells; obesity;  
 KW Berger disease; nephropathy; Schonelein-Henoch purpura; coeliac disease;  
 KW dermatitis; herpeticiformis; Crohn's disease; thalassaemia.

XX OS Homo sapiens.

XX PN US2003083462-A1.

XX PD 01-MAY-2003.

XX PF 10-DEC-2001; 2001US-00013913.

XX PR 05-JAN-1999; 99WO-US000106.

XX PR 01-SEP-1999; 99WO-US020111.

XX PR 15-SEP-1999; 99WO-US021194.

XX PR 30-NOV-1999; 99WO-US028313.

XX PR 02-DEC-1999; 99WO-US028551.

XX PR 16-DEC-1999; 99WO-US030095.

XX PR 05-JAN-2000; 2000WO-US000219.

XX PR 06-JAN-2000; 2000WO-US000376.

XX PR 11-FEB-2000; 2000WO-US003565.

XX PR 18-FEB-2000; 2000WO-US004342.

XX PR 24-FEB-2000; 2000WO-US005004.

XX PR 02-MAR-2000; 2000WO-US005841.

XX PR 15-MAR-2000; 2000WO-US006884.

XX PR 17-MAY-2000; 2000WO-US013705.

XX PR 22-MAY-2000; 2000WO-US014042.

XX PR 30-MAY-2000; 2000WO-US014941.

XX PR 02-JUN-2000; 2000WO-US015264.

XX PR 23-AUG-2000; 2000WO-US023522.

XX PR 24-AUG-2000; 2000WO-US023328.

XX PR 08-NOV-2000; 2000WO-US030952.

XX PR 10-NOV-2000; 2000WO-US030873.

XX PR 01-DEC-2000; 2000WO-US032678.

XX PR 28-FEB-2001; 2001WO-US006520.

PR 01-MAR-2001; 2001WO-US006666.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 04-SEP-2001; 2001US-00946374.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Baker KP, Botstein D, Desnovers L, Eaton DL, Ferrara N, Fong S;  
PI Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;  
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;  
PI Williams PM, Wood WI;  
XX  
XX WFI; 2003-755122/71.  
DR N-PSDB; ADD39785.  
DR  
XX  
XX  
XX New secreted and transmembrane PRO polypeptides useful for treating  
PT cancers, kidney disorders, Crohn's disease, diabetes mellitus, hyper- or  
PT hypo-insulinemia, sports injuries and arthritis.  
XX  
XX Claim 12; SEQ ID NO 194; 557pp; English.  
PS  
CC The invention relates to an isolated PRO polypeptide (secreted or  
CC transmembrane protein) having at least 80% amino acid sequence identity  
CC to an amino acid sequence chosen from 123 fully defined sequences as  
CC given in the specification (including their extracellular domains either  
CC or without their associated signal peptides). Also include are the  
CC nucleotide (NA) sequences encoding PRO, a vector comprising the PRO NA, a  
CC host cell comprising the vector, producing PRO, a chimeric molecule  
CC comprising PRO fused to a heterologous amino acid sequence, and an anti-  
CC PRO antibody. Pro is useful as molecular weight markers for protein  
CC electrophoresis and also for chromosome identification. PRO is also  
CC useful for tissue typing. PRO and PRO NA are useful as hybridisation  
CC probes for a cDNA library to isolate the full-length PRO cDNA. PRO NA is  
CC useful for generating transgenic animals or knock-out animals which are  
CC useful in development and screening useful reagents. PRO NA is also  
CC useful in gene therapy. PRO1244, PRO1286 and PRO1303 polypeptides are  
CC useful for treating cancerous tumours. PRO1250, PRO1418 and PRO1410  
CC polypeptides are useful for suppressing immune response. PRO1246  
CC polypeptide is useful for treating cardiac insufficiency disorders.  
CC PRO1246 polypeptide is also useful for treating tumours. PRO1246 and  
CC PRO1561 polypeptide are useful for stimulating calcium flux in human  
CC umbilical vein endothelial cells. PRO1265, PRO1250 and PRO1474  
CC polypeptides are useful for treating bone and/or cartilage disorders  
CC (e.g., arthritis) and wound healing. PRO1130, PRO1275 and PRO1418  
CC polypeptides are useful for treating diabetes in skeletal muscle cells  
CC and obesity. PRO1265, PRO1244 and PRO1382 polypeptides are useful for  
CC treating Berger disease or other nephropathies associated with Schonlein-  
CC Henoch purpura, coeliac disease, dermatitis, herpeticiformis or Crohn's  
CC disease. PRO1478, PRO1265, PRO1412, PRO1279, PRO1304, PRO1306, PRO1418,  
CC PRO1410 and PRO1575 are useful in treating thalassaemias. The present  
CC sequence represents a PRO protein of the invention.  
XX  
SQ Sequence 248 AA;  
Query Match 100.0%; Score 1374; DB 7; Length 248;  
Best Local Similarity 100.0%; Pred. No. 3.8e-98;  
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGLSIFLLCVLGLSQAATPKIFNGTECGRNSQVQVGLFEGTSLRCGGVLIDHRWLTA 60  
Db 1 MGLSIFLLCVLGLSQAATPKIFNGTECGRNSQVQVGLFEGTSLRCGGVLIDHRWLTA 60  
QY 61 AHCSGSRYWRLGHEHLSQLDWTETQIRHSGFSVTHPGYLCASTSHEHDLRLRLPVVR 120  
Db 61 AHCSGSRYWRLGHEHLSQLDWTETQIRHSGFSVTHPGYLCASTSHEHDLRLRLPVVR 120  
QY 121 TSSVQPLPLNDCATAGTECHVSGWGITNHPRPFPDLOCLNLSIVSHATCHGVYPGRI 180  
Db 121 TSSVQPLPLNDCATAGTECHVSGWGITNHPRPFPDLOCLNLSIVSHATCHGVYPGRI 180  
QY 181 TSNMVCAGGVPGQDACQGDGSLVCGVQLGLVSWGSGVPGQDGPVGYTYICKYVDW 240

Db 181 TSNMVCAGGVPGQDACQGDGSLVCGVQLGLVSWGSGVPGQDGPVGYTYICKYVDW 240  
QY 241 IEMIMENN 248  
Db 241 IEMIMENN 248  
RESULT 13  
ADD70232  
ID ADD70232 standard; protein; 248 AA.  
XX  
AC ADD70232;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Human secreted/transmembrane protein PRO1303.  
XX  
KW Human; secreted protein; transmembrane protein; PRO; tumour;  
KW immune response; cardiac insufficiency disorder; calcium flux;  
KW umbilical vein endothelial cell; bone disorder; cartilage disorder;  
KW arthritis; wound healing; diabetes; skeletal muscle cells; obesity;  
KW Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;  
KW dermatitis; herpeticiformis; Crohn's disease; thalassaemia.  
XX  
OS Homo sapiens.  
XX  
PN US2003054406-A1.  
XX  
PD 20-MAR-2003.  
XX  
PF 06-DEC-2001; 2001US-00006818.  
XX  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 01-SEP-1998; 98US-0098749P.  
PR 01-SEP-1998; 98US-0098750P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099536P.  
PR 09-SEP-1998; 98US-0099596P.  
PR 09-SEP-1998; 98US-0099598P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 09-SEP-1998; 98US-0099642P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099808P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 10-SEP-1998; 98US-0099815P.  
PR 15-SEP-1998; 98US-0100385P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 15-SEP-1998; 98US-0100390P.  
PR 16-SEP-1998; 98US-0100584P.  
PR 16-SEP-1998; 98US-0100627P.  
PR 16-SEP-1998; 98US-0100661P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100710P.  
PR 17-SEP-1998; 98US-0100711P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100848P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 18-SEP-1998; 98US-0101071P.  
PR 22-SEP-1998; 98US-0101279P.

PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101474P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101476P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 23-SEP-1998; 98US-0101479P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101741P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101915P.  
PR 24-SEP-1998; 98US-0101916P.  
PR 24-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102307P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102484P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-0103314P.  
PR 07-OCT-1998; 98US-0103315P.  
PR 07-OCT-1998; 98US-0103328P.  
PR 07-OCT-1998; 98US-0103395P.  
PR 07-OCT-1998; 98US-0103396P.  
PR 07-OCT-1998; 98US-0103401P.  
PR 08-OCT-1998; 98US-0103633P.  
PR 08-OCT-1998; 98US-0103678P.  
PR 08-OCT-1998; 98US-0103679P.  
PR 08-OCT-1998; 98US-0103711P.  
PR 14-OCT-1998; 98US-0104257P.  
PR 20-OCT-1998; 98US-0104987P.  
PR 20-OCT-1998; 98US-0105000P.  
PR 20-OCT-1998; 98US-0105002P.  
PR 21-OCT-1998; 98US-0105104P.  
PR 22-OCT-1998; 98US-0105169P.  
PR 22-OCT-1998; 98US-0105266P.  
PR 26-OCT-1998; 98US-0105693P.  
PR 26-OCT-1998; 98US-0105694P.  
PR 27-OCT-1998; 98US-0105807P.  
PR 27-OCT-1998; 98US-0105881P.  
PR 27-OCT-1998; 98US-0105882P.  
PR 27-OCT-1998; 98US-0106062P.  
PR 28-OCT-1998; 98US-0106023P.  
PR 28-OCT-1998; 98US-0106029P.  
PR 28-OCT-1998; 98US-0106030P.  
PR 28-OCT-1998; 98US-0106032P.  
PR 28-OCT-1998; 98US-0106033P.  
PR 28-OCT-1998; 98US-0106178P.  
PR 29-OCT-1998; 98US-0106248P.  
PR 29-OCT-1998; 98US-0106384P.  
PR 29-OCT-1998; 98US-0108500P.  
PR 30-OCT-1998; 98US-0106464P.  
PR 03-NOV-1998; 98US-0106850P.  
PR 03-NOV-1998; 98US-0106902P.  
PR 03-NOV-1998; 98US-0106905P.  
PR 03-NOV-1998; 98US-0106919P.  
PR 03-NOV-1998; 98US-0106932P.  
PR 10-NOV-1998; 98US-0106934P.  
PR 17-NOV-1998; 98US-0107783P.  
PR 17-NOV-1998; 98US-0108775P.  
PR 17-NOV-1998; 98US-0108779P.  
PR 17-NOV-1998; 98US-0108782P.  
PR 17-NOV-1998; 98US-0108788P.  
PR 17-NOV-1998; 98US-0108801P.  
PR 17-NOV-1998; 98US-0108802P.  
PR 17-NOV-1998; 98US-0108806P.

PR 17-NOV-1998; 98US-0108907P.  
PR 17-NOV-1998; 98US-0108867P.  
PR 17-NOV-1998; 98US-0108925P.  
PR 18-NOV-1998; 98US-0108848P.  
PR 18-NOV-1998; 98US-0108849P.  
PR 18-NOV-1998; 98US-0108850P.  
PR 18-NOV-1998; 98US-0108851P.  
PR 18-NOV-1998; 98US-0108852P.  
PR 18-NOV-1998; 98US-0108858P.  
PR 18-NOV-1998; 98US-0108904P.  
PR 22-DEC-1998; 98US-0113296P.  
PR 30-DEC-1998; 98US-0114223P.  
PR 05-JAN-1999; 99WO-US000106.  
PR 16-APR-1999; 99US-0129674P.  
PR 23-JUN-1999; 99US-0141037P.  
PR 20-JUL-1999; 99US-0144758P.  
PR 26-JUL-1999; 99US-0145698P.  
PR 01-SEP-1999; 99WO-US020111.  
PR 15-SEP-1999; 99WO-US021194.  
PR 29-OCT-1999; 99US-0162506P.  
PR 30-NOV-1999; 99WO-US028313.  
PR 02-DEC-1999; 99WO-US028551.  
PR 16-DEC-1999; 99WO-US030095.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 04-SEP-2001; 2001US-00946374.  
XX  
XX (GETH ) GENENTECH INC.  
PA Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;  
PI Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;  
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;  
PI Williams PM, Wood WI;  
XX WPI; 2003-708344/67.  
DR N-PSDB; ADD70231.  
XX  
XX Novel isolated PRO polypeptide useful for tissue typing, modulating  
PT biological activity of cell, as molecular weight markers in protein  
PT electrophoresis, for treating arthritis, tumor.  
XX  
PS Claim 12; SEQ ID NO 194; 549pp; English.  
XX  
CC The invention relates to an isolated PRO polypeptide (secreted or

Query Match 100.0%; Score 1374; DB 7; Length 248;  
Best Local Similarity 100.0%; Pred. No. 3.8e-98;  
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLSIFLLCVLGLSQANTPKIFNCTGCRNSQPMQVGLFEGTSLRCGGVLIDHRWLTA 60  
DB 1 MGLSIFLLCVLGLSQANTPKIFNCTGCRNSQPMQVGLFEGTSLRCGGVLIDHRWLTA 60







PR 15-SEP-1998; 98US-0100390P.  
PR 16-SEP-1998; 98US-0100584P.  
PR 16-SEP-1998; 98US-0100627P.  
PR 16-SEP-1998; 98US-0100661P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100710P.  
PR 17-SEP-1998; 98US-0100711P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100848P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 18-SEP-1998; 98US-0101071P.  
PR 22-SEP-1998; 98US-0101279P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101474P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101476P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 23-SEP-1998; 98US-0101479P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101741P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101915P.  
PR 24-SEP-1998; 98US-0101916P.  
PR 24-SEP-1998; 98US-0102207P.  
PR 23-SEP-1998; 98US-0102240P.  
PR 23-SEP-1998; 98US-0102307P.  
PR 23-SEP-1998; 98US-0102330P.  
PR 23-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102484P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 30-SEP-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-0103314P.  
PR 07-OCT-1998; 98US-0103315P.  
PR 07-OCT-1998; 98US-0103328P.  
PR 07-OCT-1998; 98US-0103395P.  
PR 07-OCT-1998; 98US-0103396P.  
PR 07-OCT-1998; 98US-0103401P.  
PR 08-OCT-1998; 98US-0103633P.  
PR 08-OCT-1998; 98US-0103678P.  
PR 08-OCT-1998; 98US-0103679P.  
PR 08-OCT-1998; 98US-0103711P.  
PR 14-OCT-1998; 98US-0104257P.  
PR 20-OCT-1998; 98US-0104987P.  
PR 20-OCT-1998; 98US-0105000P.  
PR 20-OCT-1998; 98US-0105002P.  
PR 22-OCT-1998; 98US-0105104P.  
PR 22-OCT-1998; 98US-0105169P.  
PR 22-OCT-1998; 98US-0105266P.  
PR 26-OCT-1998; 98US-0105693P.  
PR 26-OCT-1998; 98US-0105694P.  
PR 27-OCT-1998; 98US-0105807P.  
PR 27-OCT-1998; 98US-0105881P.  
PR 27-OCT-1998; 98US-0105882P.  
PR 27-OCT-1998; 98US-0106062P.  
PR 28-OCT-1998; 98US-0106023P.  
PR 28-OCT-1998; 98US-0106029P.  
PR 28-OCT-1998; 98US-0106030P.  
PR 28-OCT-1998; 98US-0106032P.  
PR 28-OCT-1998; 98US-0106033P.  
PR 28-OCT-1998; 98US-0106178P.  
PR 15-SEP-1998; 98US-0100390P.  
PR 16-SEP-1998; 98US-0100584P.  
PR 16-SEP-1998; 98US-0100627P.  
PR 16-SEP-1998; 98US-0100661P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100710P.  
PR 17-SEP-1998; 98US-0100711P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100848P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 18-SEP-1998; 98US-0101071P.  
PR 22-SEP-1998; 98US-0101279P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101474P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101476P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 23-SEP-1998; 98US-0101479P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101741P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101915P.  
PR 24-SEP-1998; 98US-0101916P.  
PR 24-SEP-1998; 98US-0102207P.  
PR 23-SEP-1998; 98US-0102240P.  
PR 23-SEP-1998; 98US-0102307P.  
PR 23-SEP-1998; 98US-0102330P.  
PR 23-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102484P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 30-SEP-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-0103314P.  
PR 07-OCT-1998; 98US-0103315P.  
PR 07-OCT-1998; 98US-0103328P.  
PR 07-OCT-1998; 98US-0103395P.  
PR 07-OCT-1998; 98US-0103396P.  
PR 07-OCT-1998; 98US-0103401P.  
PR 08-OCT-1998; 98US-0103633P.  
PR 08-OCT-1998; 98US-0103678P.  
PR 08-OCT-1998; 98US-0103679P.  
PR 08-OCT-1998; 98US-0103711P.  
PR 14-OCT-1998; 98US-0104257P.  
PR 20-OCT-1998; 98US-0104987P.  
PR 20-OCT-1998; 98US-0105000P.  
PR 20-OCT-1998; 98US-0105002P.  
PR 22-OCT-1998; 98US-0105104P.  
PR 22-OCT-1998; 98US-0105169P.  
PR 22-OCT-1998; 98US-0105266P.  
PR 26-OCT-1998; 98US-0105693P.  
PR 26-OCT-1998; 98US-0105694P.  
PR 27-OCT-1998; 98US-0105807P.  
PR 27-OCT-1998; 98US-0105881P.  
PR 27-OCT-1998; 98US-0105882P.  
PR 27-OCT-1998; 98US-0106062P.  
PR 28-OCT-1998; 98US-0106023P.  
PR 28-OCT-1998; 98US-0106029P.  
PR 28-OCT-1998; 98US-0106030P.  
PR 28-OCT-1998; 98US-0106032P.  
PR 28-OCT-1998; 98US-0106033P.  
PR 28-OCT-1998; 98US-0106178P.  
PR 29-OCT-1998; 98US-0106248P.  
PR 29-OCT-1998; 98US-0106384P.  
PR 30-OCT-1998; 98US-0106464P.  
PR 30-OCT-1998; 98US-0106566P.  
PR 03-NOV-1998; 98US-0106856P.  
PR 03-NOV-1998; 98US-0106902P.  
PR 03-NOV-1998; 98US-0106905P.  
PR 03-NOV-1998; 98US-0106919P.  
PR 03-NOV-1998; 98US-0106932P.  
PR 03-NOV-1998; 98US-0106934P.  
PR 10-NOV-1998; 98US-0107783P.  
PR 17-NOV-1998; 98US-0108775P.  
PR 17-NOV-1998; 98US-0108779P.  
PR 17-NOV-1998; 98US-0108787P.  
PR 17-NOV-1998; 98US-0108788P.  
PR 17-NOV-1998; 98US-0108801P.  
PR 17-NOV-1998; 98US-0108802P.  
PR 17-NOV-1998; 98US-0108806P.  
PR 17-NOV-1998; 98US-0108807P.  
PR 17-NOV-1998; 98US-0108867P.  
PR 17-NOV-1998; 98US-0108925P.  
PR 18-NOV-1998; 98US-0108848P.  
PR 18-NOV-1998; 98US-0108849P.  
PR 18-NOV-1998; 98US-0108850P.  
PR 18-NOV-1998; 98US-0108851P.  
PR 18-NOV-1998; 98US-0108852P.  
PR 18-NOV-1998; 98US-0108858P.  
PR 22-DEC-1998; 98US-0108904P.  
PR 30-DEC-1998; 98US-0114223P.  
PR 05-JAN-1999; 99MO-US000106.  
PR 16-APR-1999; 99US-0129674P.  
PR 23-JUN-1999; 99US-0141037P.  
PR 26-JUL-1999; 99US-0144758P.  
PR 01-SEP-1999; 99MO-US020111.  
PR 15-SEP-1999; 99MO-US021194.  
PR 23-OCT-1999; 99US-0162506P.  
PR 30-NOV-1999; 99MO-US028313.  
PR 02-DEC-1999; 99MO-US028551.  
PR 16-DEC-1999; 99MO-US030095.  
PR 06-JAN-2000; 2000MO-US000376.  
PR 11-FEB-2000; 2000MO-US003565.  
PR 18-FEB-2000; 2000MO-US004342.  
PR 24-FEB-2000; 2000MO-US005004.  
PR 02-MAR-2000; 2000MO-US005841.  
PR 15-MAR-2000; 2000MO-US006884.  
PR 17-MAY-2000; 2000MO-US013705.  
PR 22-MAY-2000; 2000MO-US014042.  
PR 30-MAY-2000; 2000MO-US014941.  
PR 02-JUN-2000; 2000MO-US015264.  
PR 23-AUG-2000; 2000MO-US023522.  
PR 24-AUG-2000; 2000MO-US023328.  
PR 08-NOV-2000; 2000MO-US030952.  
PR 10-NOV-2000; 2000MO-US030873.  
PR 03-DEC-2000; 2000MO-US032678.  
PR 28-FEB-2001; 2001MO-US0006520.  
PR 01-MAR-2001; 2001MO-US006666.  
PR 01-JUN-2001; 2001MO-US017800.  
PR 20-JUN-2001; 2001MO-US019692.  
PR 29-JUN-2001; 2001MO-US021066.  
PR 09-JUL-2001; 2001MO-US021735.  
PR 04-SEP-2001; 2001US-00946374.  
XX  
PA (GETH ) GENENTECH INC.  
XX Baker KP, Botstein D, Desnovers L, Eaton DL, Ferrara N, Fong S;  
XX Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;  
PI Pan J, Faoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;  
PI Williams PM, Wood WI;  
XX WPI; 2003-786999/74.  
DR

DR N-PSDB; ADD39308.

XX Novel isolated PRO polypeptide useful for tissue typing, modulating  
PT biological activity of cell, as molecular weight markers in protein  
PT electrophoresis, for treating arthritis, tumor.

PS Claim 12; SEQ ID NO 194; 550pp; English.

CC The invention relates to an isolated PRO polypeptide (secreted or

Query Match 100.0%; Score 1374; DB 7; Length 248;  
Best Local Similarity 100.0%; Pred. No. 3.8e-98;  
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLSIFLLLCVGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWVLT 60

Db 1 MGLSIFLLLCVGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWVLT 60

QY 61 AHCSGSRYWVRLGEHSLQOLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLPVRV 120

Db 61 AHCSGSRYWVRLGEHSLQOLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLPVRV 120

QY 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHPNPPFDLLQCLNLSIVSHATCHGVYPGRI 180

Db 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHPNPPFDLLQCLNLSIVSHATCHGVYPGRI 180

QY 181 TSNMVCAGGVPQDACCQSDSGGPLVCGGVLOGLVSWGSPCGQDGPVGYTYICKYVDW 240

Db 181 TSNMVCAGGVPQDACCQSDSGGPLVCGGVLOGLVSWGSPCGQDGPVGYTYICKYVDW 240

QY 241 IRMIMRNN 248

Db 241 IRMIMRNN 248

Search completed: June 16, 2004, 20:32:43  
Job time : 63 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 20:30:31 ; Search time 20 Seconds  
(without alignments)  
1192.775 Million cell updates/sec

Title: US-10-015-385A-194

Perfect score: 1374  
Sequence: 1 MGSLIFLLCVLGLSQAATP.....GVYYICKYVDIMRMNRN 248

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: Pirl1.\*  
2: Pirl2.\*  
3: Pirl3.\*  
4: Pirl4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	622.5	45.3	260	2 I56559	neutropsin - mouse
2	569.5	41.4	248	2 S55066	trypsin EC 3.4.21
3	567	41.3	248	2 S55067	trypsin EC 3.4.21
4	561	40.8	231	1 TRPGTR	trypsin EC 3.4.21
5	555	40.4	243	2 A35871	trypsin EC 3.4.21
6	555	40.4	253	2 A53968	serine proteinase
7	553	40.2	246	2 B25528	trypsin EC 3.4.21
8	552	40.2	229	1 TREOTR	trypsin EC 3.4.21
9	550	40.0	247	2 S13813	trypsin EC 3.4.21
10	546	39.7	247	1 TRDG	trypsin EC 3.4.21
11	545	39.7	246	1 TRDGC	trypsin EC 3.4.21
12	544.5	39.6	247	2 A27547	trypsin EC 3.4.21
13	543	39.5	246	1 TRRT1	trypsin EC 3.4.21
14	540	39.3	238	2 S13779	trypsin EC 3.4.21
15	533.5	38.8	247	2 S54494	trypsin EC 3.4.21
16	532	38.7	246	1 TRRT2	trypsin EC 3.4.21
17	527.5	38.4	261	2 A31136	tissue kallikrein
18	524	38.1	232	1 KOPG	tissue kallikrein
19	520.5	37.9	261	2 A25606	tissue kallikrein
20	519	37.8	246	2 JQ1471	trypsin EC 3.4.21
21	518	37.7	246	2 JQ1472	trypsin EC 3.4.21
22	518	37.7	261	2 A34079	tissue kallikrein
23	513.5	37.4	231	2 S31778	trypsin EC 3.4.21
24	513.5	37.4	239	2 A27207	tissue kallikrein
25	511.5	37.2	261	2 S45303	tissue kallikrein
26	510.5	37.0	261	1 NGMSG	7S nerve growth fa
27	508	37.0	242	2 S31775	trypsin EC 3.4.21
28	508	37.0	242	2 S31776	trypsin EC 3.4.21
29	503	36.6	260	2 A37938	tissue kallikrein

30 503 36.6 263 2 S15686 tissue kallikrein  
31 502 36.5 259 2 B31136 tissue kallikrein  
32 501 36.5 247 1 B25852 trypsin EC 3.4.21  
33 499.5 36.4 241 2 S39048 trypsin EC 3.4.21  
34 496.5 36.1 261 1 TRMSM5 tissue kallikrein  
35 496 36.1 242 2 S49489 trypsin EC 3.4.21  
36 495.5 36.1 257 2 S33772 tissue kallikrein  
37 495.5 36.1 261 2 S01971 tissue kallikrein  
38 494.5 36.0 247 1 A25852 trypsin EC 3.4.21  
39 493 35.9 262 1 KQHU tissue kallikrein  
40 491.5 35.8 240 2 S39047 trypsin EC 3.4.21  
41 491 35.7 256 1 NGMSA 7S nerve growth fa  
42 490.5 35.7 229 1 TRDFS trypsin EC 3.4.21  
43 490.5 35.7 261 2 A41020 tissue kallikrein  
44 489 35.6 244 2 A44284 tissue kallikrein  
45 485.5 35.3 261 2 JE0236 tissue kallikrein

## ALIGNMENTS

RESULT 1  
I56559  
neutropsin - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 21-Jul-2000  
C:Accession: I56559  
R:Chen, Z.L.; Yoshida, S.; Kato, K.; Momota, Y.; Suzuki, J.; Tanaka, T.; Ito, J.; Ni J. Neurosci. 15, 5088-5097, 1995  
A:Title: Expression and activity-dependent changes of a novel limbic-serine protease  
A:Reference number: I56559; MUID:95348817; PMID:7623137  
A:Accession: I56559  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-260 <RES>  
A:Cross-references: GB:D30785; NID:G1648847; PIDN:BAA06451.1; PID:G1020091  
C:Superfamily: trypsin; trypsin homology  
F:33-252/Domain: trypsin homology <TRY>

Query Match 45.3%; Score 622.5; DB 2; Length 260;  
Best Local Similarity 49.0%; Pred. No. 1.4e-43;  
Matches 121; Conservative 28; Mismatches 93; Indels 5; Gaps 3;

Qy 5 IFLLLCV---LGLSQAATPKIFNGTEGRNSQPMQVGLFEGTSRCCGVLIDHWLTAA 61  
Db 13 ILLLFMGAGAGLTRAQGSKILEGECIPHSPQMAALFQGERLICGGVLGDRWLTAA 72  
Qy 62 HCSGRYVRLGEHSLQLDWTEQIRHSGFSVTHPGVLGAS-TSHEHDLRLRLRPVRV 120  
Db 73 HCKQKYSVRLGDHSLQSRDQPEIQVAQSIQHPCTVNSNPEDSHDMLIRLQNSANL 132  
Qy 121 TSSVOPLPLPNDCTAGTECHVSGWGTNHPNPFDPDLQCLNLSIVSHATCHGVYFGR 180  
Db 133 GDKVKPQLANLCPKVQKQCIISGWGTWTSQENFPNTLNCVAKVIYSQNKCEAYPGKI 192  
Qy 181 TSNMVCAGVPGQDACCGDSGGLVCGVLOGLVSWSGVPCGGQDGPVYTYICKYVDW 240  
Db 193 TEGMVCAGSSNGADTCQDSGGPLVCDMLQGITWSGS-DPCGRPEKPGVYTKICRYTTW 251  
Qy 241 IRIMLRN 247  
Db 252 IKKTMND 258

RESULT 2  
S55066  
trypsin (EC 3.4.21.4) II precursor, pancreatic (clone 2-P29) - chicken  
N:Alternate names: trypsinogen II  
C:Species: Gallus gallus (chicken)  
C:Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 22-Jun-1999  
C:Accession: S55066; S72347  
R:Wang, K.; Gan, L.; Lee, I.; Hood, L.  
Biochem. J. 307, 471-479, 1995



```
F;60,62,65,70/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted
Query Match 40.4%; Score 561; DB 1; Length 231;
Best Local Similarity 46.7%; Pred. No. 1.3e-38;
Matches 107; Conservative 39; Mismatches 77; Indels 6; Gaps 4;

QY 21 KTFNGTEGCRNSQPMQVGLFEGTSLRCGGVLIDHRWLTAHCSSRYWRLGHSLSQL 80
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 8 KIVGYTCAANSIPYQVSLNSGSHF--CGGLSLNSQWVSAARHCYKSRIQVRLGEHNDVL 66
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 81 DWTEQIRHSGFVTHPGYLGASTSHEHDLRLRLRPVVTSSVQPLPNDCAATGTC 140
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 67 EGNEQFINAKIITHPNFNG--NTLNDIMLKSPATINSRVATVSLPRCAAGTGC 124
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 141 HVSNGWITNHPNPPDLLQCLNLSIVSHATGCHGVYVPGRTSNMVCAGGVP--GQDACQGD 199
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 125 LLSGNGTKSSGSSYPSLLQCLKAPVLSDSCKSSYPGQITGNMVCVFLGEGKDCSQGD 184
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 200 SGGPLVCGVGLQGLVSWGSGVPCGGQIGPVYTYCKYVDWIRMNRNN 248
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 185 SGGPVVNCQLOGIVSWGY--GCAQKNKPGVYTKVCNVYNNWIIQOTIAAN 231
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
A35871
trypsin (EC 3.4.21.4) precursor, pancreatic - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 04-Mar-2000
C:Accession: A35871; S12117
R:Shi, Y.B.; Brown, D.D.
Genes Dev. 4, 1107-1115, 1990
A:Title: Developmental and thyroid hormone-dependent regulation of pancreatic genes in X
A:Reference number: A35871; MUID:91007255; PMID:2210372
A:Accession: A35871
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-243 <SHI>
A:Cross-references: EMBL:X53458; NID:965162; PIDN:CAA37538.1; PID:965163
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; protein digestion; serine proteinase
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-20/Domain: activation peptide #status predicted <APT>
F:21-236/Domain: trypsin homology <TRY>
F:26-243/Product: trypsin I #status predicted <MAT>
F:27-157,45-61,129-230,136-203,168-182/Disulfide bonds: #status predicted
F:60,104,197/Active site: His, Asp, Ser #status predicted

Query Match 40.4%; Score 555; DB 2; Length 243;
Best Local Similarity 45.7%; Pred. No. 4.3e-38;
Matches 113; Conservative 42; Mismatches 82; Indels 10; Gaps 6;

QY 6 FLLLCVLGLSQAAT---PKIFNGTEGCRNSQPMQVGLFEGTSLRCGGVLIDHRWLTAH 62
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 FLLLCVL-LGAAAFDDDKIIGCATCAKSSVPVIVSLNSGYHF--CGGLSLTNQWVSAAH 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 63 CSSSRWRLGHSLSQLDWTQIRHSGFVTHPGYLGASTSHEHDLRLRLRPVVT 122
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 CYKASIQVRLGHNIALSEGTQFISSSKVIRSGY--NSYTLNDIMLIKLSPPASLNA 118
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 123 SVQPLPNDCAATGTECHVSGWGTINHPNPPDLLQCLNLSIVSHATGCHGVYVPGRTS 182
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 119 AVNTVPLPSCSAAGTSCILSHGNTLSNGSNVFDLLQCLNAPILINAQCNAYPGEITA 178
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 183 NMVCAGGVP--GQDACQGDSSGGPLVCGVGLQGLVSWGSGVPCGGQIGPVYTYCKYVDWI 241
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 NMICVGMGEGKDCSQDGGPVVNCQLOGVSWGY--GCAMRNYPGVYTKVCNVNNAWI 236
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 242 RMIMNRNN 248
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 237 QNTIAAN 243
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
```

```
A53968
serine proteinase SCCE precursor - human
N:Alternate names: stratum corneum chymotryptic enzyme
C:Species: Homo sapiens (man)
C>Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 22-Jun-1999
C:Accession: A53968
R:Hansson, L.; Stromqvist, M.; Baackman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud
J. Biol. Chem. 269, 19420-19426, 1994
A:Title: Cloning, expression, and characterization of stratum corneum chymotryptic e
A:Reference number: A53968; MUID:94308225; PMID:8034709
A:Accession: A53968
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-253 <HAN>
A:Cross-references: GB:L33404; NID:9521214; PIDN:AAC37551.1; PID:9532504
C:Genetics:
A:Gene: GDB:PRSS6; SCCE
A:Cross-references: GDB:377730
A:Map position: 7q35-7q35
C:Superfamily: trypsin; trypsin homology
F:30-245/Domain: trypsin homology <TRY>

Query Match 40.4%; Score 585; DB 2; Length 253;
Best Local Similarity 44.6%; Pred. No. 4.5e-38;
Matches 115; Conservative 37; Mismatches 90; Indels 16; Gaps 6;

QY 1 MGLSIFLLLCVLGLS-----QAATPKIFNGTEGCRNSQPMQVGLFEGTSLRCGGVLI 52
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MARSULLPLQIILLLSALETAGEEAQQGKIIDGAPCARGHPWQVALLSGNQLHCCGVILV 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 53 DHRWVLTAAHCSSRYWRLGHSLSQLDWTQIRHSGFVTHPGYLGASTSHEHDLRL 112
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 NERWVLTAAHCXKNEYVHLGSDTLG--DRRAQRIKASKSFRHPGY--STQTHVNDLMV 116
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 113 RLRLPVRTSSVQPLPNDCAATGTECHVSGWGTINHPNPPDLLQCLNLSIVSHATC 172
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 117 KLSNQARLSSMVKKVRLPSRCPPGTTCTVSCWGTTTSPDVTFFPSDLMCVDVKLLSPQC 176
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 173 HGVPGRITSNMVCAGGVP--QDACQGDSSGGPLVCGVGLQGLVSWGSGVPCGGQIGPV 230
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 177 TKVYKDLLENMLCA-GIPDSKKNACNGSGGLVCRGTLOGVLVSWGTF--PCGQPNPGV 234
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 231 YTVICKYVDWIRMNRNN 248
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 235 YTVQCKFTWINDTKKH 252
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
B25528
trypsin (EC 3.4.21.4) precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 22-Jun-1999
C:Accession: B25528
R:Stevenson, B.J.; Hagenbuechle, O.; Wellauer, P.K.
Nucleic Acids Res. 14, 8307-8330, 1986
A:Title: Sequence organisation and transcriptional regulation of the mouse elastase
A:Reference number: A93646; MUID:87066713; PMID:3641189
A:Accession: B25528
A:Molecule type: mRNA
A:Residues: 1-246 <STE>
A:Cross-references: GB:X04574; NID:954918; PIDN:CAA28243.1; PID:954919
C:Superfamily: trypsin; trypsin homology
C:Keywords: calcium binding; hydrolase; protein digestion; serine proteinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-239/Product: trypsin #status predicted <MAT>
F:24-239/Domain: trypsin homology <TRY>
F:30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
F:63,107,200/Active site: His, Asp, Ser #status predicted
F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 40.2%; Score 553; DB 2; Length 246;
Best Local Similarity 44.7%; Pred. No. 6.3e-38;
Matches 113; Conservative 45; Mismatches 81; Indels 14; Gaps 7;
```

Qy	3	LSIFLLLCVLGLSQA----	ATPKTFNGTECGRNSQPMOVL	FGTSLRCGGVLIDHRWTL	58
Db	1	MSALITLALVGAAPFVDDDK	LVGGYTCRESSVPYQVSLNAGYHF	-CGGSLINDQWV	59
Qy	59	TAAHCGSGRYWYRLGEHLSLQ	LDWTEQIRHSGFVTHPGYLGAST	SGHEHDLRLRLRLP	118
Db	60	SAACHVKYRIQVRLGEHNIN	LEGNEQFVDSAKIIRHPNY	--NSWTLDNDIMLIKASPV	117
Qy	119	RYTSSVQPIPLPNDCATAGT	CHVSGWG--ITNHPRAPFPDL	LQCLNLSIVSHATCGVY	176
Db	118	TUNARVASYPLESSCAPAGT	QCLISGWNTLSNGVN--PDLLQ	CVDPAPVLPOADCEASY	175
Qy	177	PGRITSNMVCAGVP--GODAC	CGDSGGPLVCGGYLQGLVSWG	SVSGPGQDGIQGVYTYIC	235
Db	176	PGDITNNMLCVFSLGGKDS	CQDSGGFPVNCGELQIVSWG	Y--GCAQPDAPGVYTKVC	233
Qy	236	KYVDWIRMTIMRNN	248		
Db	234	NYVDWIONTIADN	246		

RESULT 8  
TRYPSIN  
trypsin (EC 3.4.21.4) precursor - bovine  
N;Contains: trypsinogen  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 24-Apr-1984 #sequence revision 28-Feb-1986 #text\_change 18-Jul-1997  
C;Accession: A90164; A00946; S08774  
R;Mikes, O.; Holsayovsky, V.; Tomasek, V.; Sorm, F.  
Biochem. Biophys. Res. Commun. 24, 346-352, 1966  
A;Title: Covalent structure of bovine trypsinogen. The position of the remaining amides.  
A;Reference number: A90164; MUID:67168848; PMID:5967094  
A;Accession: A90164  
A;Molecule type: protein  
A;Residues: 1-57, 'Q', 59-67, 'Q', 69-150, 'N', 152-176, 'N', 178-229 <MKT>  
R;Hartley, B.S.  
Philos. Trans. R. Soc. Lond. B257, 77-87, 1970  
A;Reference number: A93755  
A;Contents: annotation; revisions  
R;Titani, K.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.  
Biochemistry 14, 1389-1366, 1975  
A;Title: Amino acid sequence of dogfish trypsin.  
A;Reference number: A00950; MUID:75146445; PMID:1092332  
A;Contents: annotation; revisions  
A;Note: the sequence agrees with that shown  
R;Bode, W.; Schwager, P.  
J. Mol. Biol. 98, 693-717, 1975  
A;Title: The refined crystal structure of bovine beta-trypsin at 1.8 angstrom resolution  
A;Reference number: A92954; MUID:76072097; PMID:512  
A;Contents: annotation; X-ray crystallography; binding sites for calcium, substrate, and  
C;Comment: Trypsinogen is synthesized in the acinar cells of the pancreas.  
C;Comment: Autocatalytic cleavage after Lys-6 leads to beta-trypsin by releasing a terminal  
s pseudotrypsin. A cleavage may also occur after Arg-105.  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen  
F;1-229/Product: trypsinogen #status experimental <zym>  
F;1-6/Domain: activation peptide #status experimental <zym>  
F;7-222/Domain: trypsin homology <TRY>  
F;7-131, 132-229/Product: alpha-trypsin #status experimental <MPT>  
F;6-7/Cleavage site: Lys-Ile (enteropeptidase) #status experimental  
F;13-143, 31-47, 115-216, 222-189, 154-168, 179-203/Disulfide bonds: #status experimental  
F;46, 90, 183/Active site: His, Asp, Ser #status experimental  
F;58, 60, 63, 68/Binding site: calcium (Glu, Asn, Val, Glu) #status experimental  
F;131-132/Cleavage site: Lys-Ser (autolytic) #status experimental

Query Match	40.2%	Score 552;	DB 1;	Length 229;
Best Local Similarity	46.3%;	Pred. No. 7.1e-38;		
Matches 106;	Conservative 40;	Mismatches 77;	Indels 6;	Gaps 4;
Qy	21	KIFNGTTCGRNSQPWQGLEFEGTSLRCGGVLIDHRWLTAAHSCSGRYWVRLGHEHLSQL	80	
Db	6	KIVGGYTTCGANTVPYQVSLNSGYHF-CGGSLINSQWVVSAAHCYKSGIOVRLGEDINNV	64	

QY	81	DWTEQIRHSGFSYTHPGVLGASNSHEHDLRLRLRLPVRVYSSVQPLPLNDCAAGATGEC	140
Db	65	EGNEQFISAKSIVHPSY--NSNTLNDNDMLIKSAASNSRVASISLPTSCASAGTQC	122
QY	141	HVSGKGIETHNRPFPDLLQCLNLSIVSHATCGHYVPGRIITSNMVCAGVP-GQDACQGD	199
Db	123	LISGWNGTKSSGTSYPDLVKCLKAPIILSDSKSAPGQITSNMPCAGYLEGGKDCSCGD	182
QY	200	SGGPLVCGGLQGLVMSGVSPGQGDGIPGYTVYICKYVDWIRIMRN	248
Db	183	SGGPVVCSGKLGQIVSNGS--GCAQKNKPGVYTKVCNVYSWIKOTIASN	229

```

RESULT 9
S13813
  trypsin (EC 3.4.21.4) - bovine
C/Species: Bos primigenius taurus (cattle)
C/Date: 02-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
C/Accession: S13813
R/Le Huorou, I.; Wicker, C.; Guilloteau, P.; Toullec, R.; Puigserver, A.
Eur. J. Biochem. 193; 767-773; 1990
A/Title: Isolation and nucleotide sequence of cDNA clone for bovine pancreatic anion
A/Reference number: S13813; MUID:91065383; PMID:1701147
A/Accession: S13813
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-247 <HUE>
A/Cross-references: EMBL:X54703; NID:g829; PIDN:CAA38513.1; PID:g830
C/Superfamily: trypsin; trypsin homology
C/Keywords: hydrolase; protein digestion; serine proteinase
F/24-239/domain: trypsin homology <TRY>
F/63,107,200/Active site: His, Asp, Ser #status predicted

```

Query Match	40.0%;	Score	550;	DB	2;	Length	247;
Best local Similarity	45.7%;	Pred. No.	1.1e-37;				
Matches	113;	Conservative	40;	Mismatches	84;	Indels	10;
Gaps	5;						

  

QY	7	LLLCVLGLSQA----	ATPKIENGTECGRNSQSPQWGLFEGTSLRCGGVLIDHRWVLTAAH	62
Db	5	LILAFVGAAVAFPSDDDDKIVGGYTCRAENSVPYQVSLNAGYHF-CGGSLINDQWVVSAAH	63	
QY	63	CGSGRYVVRIGEHLSOLDWTEQTRHSGFSVTHCYLGASTSHEDHLRLRLRLPVRVTS	122	
Db	64	CYQVHIQVRLEJYINIDVLEGGQIDASKIIRHPKY--SSWFLDNILILKILSTFAVINA	121	
QY	123	SVQPLPLFNDCATAGTECHVSGWGITNHRPNFPDLLQLCLNLSIVSHATCHGVYGPRIITS	182	
Db	122	RVSTFLLLFSACASAGTECLISGWNGLTSSGGVNPYDQLQLVAPLLSHADCEASYPGQITN	181	
QY	183	NWVCAGGVP-GODACQGDGGPLVCGGVQLGVSMGVSVPQCGDQIPGVYTVYICKYVDMI	241	
Db	182	NMICAGFLEGGSCQGDGGFVACNGQLQGVISWGY--GCAQKQKPGVYTKVCNVYDVI	239	
QY	242	RMIMRNN	248	
		:	:	
		:	:	
Db	240	QETIAAN	246	
		:	:	
		:	:	

```

RESULT 10
TRDG
trypsin (EC 3.4.21.4) precursor, anionic - dog
N/Alternate names: cationic trypsinogen
C/Species: Canis lupus familiaris (dog)
C/Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
C/Accession: A26273
R/Pinsky, S.D.; LaForge, K.S.; Scheele, G.
Mol. Cell. Biol. 5, 2669-2676, 1985
A/Title: Differential regulation of trypsinogen mRNA translation: full-length mRNA &
A/Reference number: A26273; MUID:85284628; PMID:3841794
A/Accession: A26273
A/Molecule type: mRNA
A/Residues: 1-247 <PIN>

```

QY	3	LSLFLLLCVLGSQA-----ATPKIFNGTECGRNSQPWQVLPEGTSRLRCGGVLIDHRWVL	58
Db	1	MKTFFIALLGATVAEPIDDDDKIVGGYTCRSNVPEQVSLNSGYHF-CGGSLINSQWVV	59
QY	59	TTAAHCGSGRWVRLGSHSLSQLDWTQIIRHSGFSVTHPGYLGASTSHEHDLRLRLRLP	118
Db	60	SAANCYKSRIOVRLGYNIAVSSGGROFNAAAKIIRPHY-NANTY-DDNDIMLIKLSSPA	117

trypsin [EC 3.4.21.4] I precursor - rat  
N;Alternate names: trypsinogen I  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 24-Sep-1999  
C;Accession: B22657; A06948  
R;Crack, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.  
J. Biol. Chem. 259, 14255-14264, 1984  
A;Title: Structure of two related rat pancreatic trypsin genes.  
A;Reference number: A22657; PMID:85054880; PMID:6094547  
A;Accession: B22657  
A;Molecule type: DNA  
A;Residues: 1-246 <CRA>

Matches 108; Conservative 37; Mismatches 78; Indels 6; Gaps 4;

0  
0  
0  
0  
  
4  
4  
  
. .  
0  
III  
4  
4  
0  
  
0  
0  
0

Search completed: June 16, 2004, 20:34:43  
Job time : 21 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 20:30:01 ; Search time 45 Seconds  
(without alignments)  
1738.856 Million cell updates/sec

Title: US-10-015-385A-194  
Perfect score: 1374  
Sequence: 1 MGSLIFLLCVGLSQAATP.....GVYTYCKYVDWIMRMNN 248

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

1: sp\_archea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mmc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rviro:\*

16: sp\_bacteriap:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	952.5	69.3	234	11 Q9CV76	Q9cv76 mus musculu
2	627.5	45.7	260	4 Q8IW69	Q8iw69 homo sapien
3	608	44.3	254	11 Q8CGR4	Q8cgr4 mus musculu
4	601	43.7	275	4 Q8IXD7	Q8ixd7 homo sapien
5	597.5	43.5	249	11 Q9QYNA	Q9qyn4 mus musculu
6	597.5	43.5	276	11 Q9QYN3	Q9qyn3 m hippostas
7	595	43.3	255	4 Q9GRQ0	Q9grq0 homo sapien
8	588.5	42.8	293	11 Q9D140	Q9d140 mus musculu
9	588	42.8	276	11 Q8CGR6	Q8cgr6 mus musculu
10	586	42.6	250	11 Q8CGR5	Q8cgr5 mus musculu
11	578	42.1	242	11 Q8OVSA	Q8ovs4 mus musculu
12	563	41.0	248	13 Q7SZT1	Q7szt1 xenopus lae
13	555	40.4	243	13 Q7SZ06	Q7sz06 xenopus lae
14	553.5	40.3	246	11 Q88301	Q88301 mus musculu
15	553.5	40.3	253	11 Q91Y82	Q91y82 mus musculu
16	550.5	40.1	247	11 Q9D7Y7	Q9d7y7 mus musculu

17	549.5	40.0	247	11 Q9CPN9	Q9cpn9 mus musculu
18	546.5	39.8	251	11 Q54854	Q54854 rattus norv
19	544	39.6	253	4 Q8N5N9	Q8n5n9 homo sapien
20	541.5	39.4	247	11 Q9CPN7	Q9cpn7 mus musculu
21	538.5	39.2	249	11 Q91VE3	Q91ve3 mus musculu
22	536	39.0	246	11 Q9ROT7	Q9rot7 mus musculu
23	536	39.0	246	11 Q7TT42	Q7tt42 mus musculu
24	533	38.8	246	11 Q9ZLR9	Q9zlr9 mus musculu
25	532	38.7	246	11 Q9QDK9	Q9qdk9 mus musculu
26	525	38.2	240	13 Q98TH0	Q98th0 engraulis j
27	522	38.0	244	13 Q8QGW3	Q8qgw3 anguilla ja
28	514.5	37.4	242	13 Q7TIR8	Q7tir8 pangasius h
29	514	37.4	239	11 Q63275	Q63275 rattus norv
30	513.5	37.4	243	13 Q8AV83	Q8av83 brachydanio
31	513.5	37.4	261	6 Q9NIQ1	Q9niqi sagunius oe
32	513.5	37.4	278	11 Q99M20	Q99m20 mus musculu
33	511.5	37.2	261	6 Q29474	Q29474 canis famil
34	510.5	37.2	238	13 Q9W7Q6	Q9w7q6 paralichthy
35	509.5	37.1	222	13 Q8AV11	Q8avl1 oncorhynch
36	509.5	37.1	235	11 Q63274	Q63274 rattus norv
37	508	37.0	242	13 Q9W7Q7	Q9w7q7 paralichthy
38	507	36.9	244	13 Q42159	Q42159 petromyzon
39	506	36.8	245	13 Q42160	Q42160 petromyzon
40	505	36.8	242	13 Q7SX90	Q7sx90 brachydanio
41	505	36.8	249	13 Q92046	Q92046 dissoatichu
42	504.5	36.7	269	4 Q8IU55	Q8iu55 homo sapien
43	503	36.6	241	13 Q98TG9	Q98tg9 engraulis j
44	502	36.5	249	13 Q9W6K0	Q9w6k0 notothenia
45	501.5	36.5	237	13 Q91515	Q91515 fugu rubrip

## ALIGNMENTS

### RESULT 1

Q9CV76	PRELIMINARY;	PRT;	234 AA.
ID	Q9CV76	PRELIMINARY;	PRT;
AC	Q9CV76;		
DT	01-JUN-2001 (TRENBLrel. 17, Created)		
DT	01-JUN-2001 (TRENBLrel. 17, Last sequence update)		
DE	01-OCT-2003 (TRENBLrel. 25, Last annotation update)		
DE	2310008B01RIK protein (Fragment).		
GN	2310008B01RIK.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Tongue;		
RX	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,		
RA	Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustineich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerke P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,		
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,		
RA	Hayashizaki Y.;		
RT	Functional annotation of a full-length mouse cDNA collection.;		
RL	Nature 409:685-690(2001).		
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.		
DR	EMBL; AK009217; BAB26143.1; -.		

```
DR HSP; P00763; IDPO.
DR MEROPS; S01.020; -.
DR MGD; MGI:1916761; 2310008B01Rik.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON TER 1
SQ SEQUENCE 234 AA; 25888 MW; 6D81E609EDD39110 CRC64;

Query Match 69.3%; Score 952.5; DB 11; Length 234;
Best Local Similarity 70.6%; Pred. No. 4.1e-84;
Matches 166; Conservative 27; Mismatches 41; Indels 1; Gaps 1;

QY 14 LSQAATPKIENGTECGNSQPMQVGLFEGTSLRCGGVLIHRWVLTAAHCSGRYVRLG 73
Db 1 LSQAADREKIYNGVECVKNSQPMQVGLFEGTSLRCGGVLIHRWVLTAAHCR-DKYVYVRLG 59
QY 74 EHSLSQDWTQEIHRHSGFSVTHPGYLGASTSHEHDLRLRLPLRVTSVQVPLPND 133
Db 60 EHSITKLDWTQELRHFTSITHPSYQAYQNHEDLRLRLNPIHLTRAVRPVLPSSC 119
QY 134 ATAGTECHVSGWGITNHRNPFDPDLQLCLNSIVSHATCHGVYVGRITSNMVCAGGVP 193
Db 120 VTTGAMCHVSGWGITNHRNPFDPDLQLCLNSIVSHATCHGVYVGRITSNMVCAGGVP 179
QY 194 DACQSGGGLVCGVQLGVLSVSGVPCGQDGPVGYTYICKYVDWIRIMINN 248
Db 180 DACQSGGGLVCGVQLGVLSVSGVPCGQDGPVGYTYICKYVDWIRIVIRN 234

RESULT 2
Q8IW69 Q8IW69 PRELIMINARY; PRT; 260 AA.
AC Q8IW69;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Srausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC040887; AA040887.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein.
SQ SEQUENCE 260 AA; 28090 MW; EF5934EB96295660 CRC64;

Query Match 44.3%; Score 608; DB 11; Length 254;
Best Local Similarity 49.0%; Pred. No. 1e-50;
Matches 124; Conservative 30; Mismatches 77; Indels 22; Gaps 7;

QY 5 IFLLLCVLGLSQAAT--PKIFNGTECGNSQPMQVGLFEGTSLRCGGVLIHRWVLTAAH 62
Db 1 MWLLAFVLLVSAAQDGVLEGECEVPHSQPMQVGLFEGTSLRCGGVLIHRWVLTAAH 60

Query Match 45.7%; Score 627.5; DB 4; Length 260;
Best Local Similarity 50.2%; Pred. No. 1.4e-52;
Matches 123; Conservative 24; Mismatches 93; Indels 5; Gaps 3;

QY 5 IFLLLCVLGLSQAATPKIENGTECGNSQPMQVGLFEGTSLRCGGVLIHRWVLTAA 61
Db 13 MFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPMQVGLFEGTSLRCGGVLIHRWVLTAA 72
QY 62 HCSGSRYWVRLGHSLSQLDWTQEIHRHSGFSVTHPGYLGAST-SHEHDLRLRLPLRV 120
Db 73 HCKPKYTVRLGHSLSQLDWTQEIHRHSGFSVTHPGYLGAST-SHEHDLRLRLPLRV 132
QY 121 TSSVQVPLPNDCATAGTECHVSGWGITNHRNPFDPDLQLCLNSIVSHATCHGVYVGR 180
Db 133 GSKVKPISLADHCTQPGQKCTVSGWGITNHRNPFDPDLQLCLNSIVSHATCHGVYVGR 192
QY 181 TSNMVCAGGVPQDADCGSGGLVCGVQLGVLSVSGVPCGQDGPVGYTYICKYVDW 240
Db 193 TDVMVCAGSGKADTCQSGGGLVCGVQLGVLSVSGVPCGQDGPVGYTYICKYVDW 251
QY 241 IRIM 245
Db 252 IKXII 256

RESULT 3
Q8CGR4 Q8CGR4 PRELIMINARY; PRT; 254 AA.
AC Q8CGR4;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Prostins.
GN KKL15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22325484; PubMed=12437987;
RA Olsson A.Y.; Lundwall A.;
RT "organization and evolution of the glandular kallikrein locus in Mus musculus.";
RL Biochem. Biophys. Res. Commun. 299:305-311 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Adams M.; Mural R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY152434; AN78422.1; -.
DR MGD; MGI:2447533; Kkl15.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 254 AA; 28042 MW; AA9E38BEBD001861 CRC64;

Query Match 44.3%; Score 608; DB 11; Length 254;
Best Local Similarity 49.0%; Pred. No. 1e-50;
Matches 124; Conservative 30; Mismatches 77; Indels 22; Gaps 7;

QY 5 IFLLLCVLGLSQAAT--PKIFNGTECGNSQPMQVGLFEGTSLRCGGVLIHRWVLTAAH 62
Db 1 MWLLAFVLLVSAAQDGVLEGECEVPHSQPMQVGLFEGTSLRCGGVLIHRWVLTAAH 60
```

```
QY 63 CSGSRVYVRLGHSLSQDWTQIRHSGFSVTHPGYLGASTSHEDLRLRLRPVTVTS 122
Db 61 CQTRFMRVRLGHNLRKFDGPEQLRSVSRIPHPGY--EARTHHRDLMLELRFKPARLTA 118
QY 123 SVQPLPLPNDCATAGTCHVSGWGITNHPNP-----FPDLQLCLNLSIVGSHAT 171
Db 119 YVRFPVALPRCPDIGEDCVVSGWGLLS-DNNPGATGSKSHVRUPDLTHCANISIISEAS 177
QY 172 CHGVYFGRITSNMVCY---GGVPQDQACQDGGGGLVCGGVLQGLVSGVSGVPGCGDGP 228
Db 178 CNKDYFGRVLPMTWCAGVEGG--GTDSCGDSGGGPIVCGGALQGLVSGGV-PCDTITKP 234
QY 229 GVTYICKYVDWI 241
Db 235 GVTYKVCYLEWI 247
RESULT 4
ID Q8IXD7 PRELIMINARY; PRT; 275 AA.
AC Q8IXD7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE variant form hippostasin/KLK11.
GN KLK11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Nakamura T., Mitsui S., Miki T., Yamaguchi N.;
RT "Molecular cloning and expression of a variant form of
RT hippostasin/KLK11 in prostate.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB078780; BAC34105.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 275 AA; 30165 MW; 257A42B28F40E2C4 CRC64;
```

```
Query Match 43.7%; Score 601; DB 4; Length 275;
Best Local Similarity 44.2%; Pred. No. 5.5e-50;
Matches 121; Conservative 35; Mismatches 88; Indels 30; Gaps 5;
QY 3 LSIFLLCVLGSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWLTAH 62
Db 4 LQILALATGLVGGET-RIIKGFCKPHSQPWQAALFEKRLCCGATLIAPRWLLTAH 62
QY 63 CSG-----SRYVRLGHSLSQDWTQIRHSGFSVTHPG 97
Db 63 CLKPWVSLTSPHVSPLSSNVCVLSHLSRYIVHLGQHLQKEGCEQRTATFPPHP 122
QY 98 YLGA--STSHEDLRLRLRPVTVTSVQPLPNDCATAGTCHVSGWGITNHPNP 155
Db 123 FNNSLPNKDRNDIMLVKASPVSIWVAPLTLSSRCVTAGTSLISGSGTSSPQLRL 182
QY 156 PDLLQCLNLSIVSHATCHGVYFGRITSNMVCYAG-GVPQDQACQDGGGGLVCGVLQGLV 214
Db 183 PHTLRANITITIEHQKCNAYFPGNITDTWVCASVQEGGKDCQDGGGGLVCGVNSQIQII 242
```

```
QY 215 SWGSGVPGQDGIPIGVYTYICKYVDWIRMNRN 248
Db 243 SWGQ-DPCAIRKPGVYTKVCKYVDWIQETMKN 275
RESULT 5
QYQYN4
ID QYQYN4 PRELIMINARY; PRT; 249 AA.
AC QYQYN4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hippostasin.
GN PRSS20.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Yamaguchi N., Mitsui S.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Mitsui S., Yamaguchi N.;
RT "cDNA cloning of a novel brain serine protease, Hippostasin.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB016226; BAA88825.1; -.
DR HSSP; P00763; LDPO.
DR MEROPS; S01.257; -.
DR MGD; MGI:192977; Prss20.
DR GO; GO:0005576; C:extracellular; IDA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 249 AA; 27604 MW; F9FF9CB457D727D5 CRC64;
```

```
Query Match 43.5%; Score 597.5; DB 11; Length 249;
Best Local Similarity 47.0%; Pred. No. 1.1e-49;
Matches 118; Conservative 33; Mismatches 95; Indels 5; Gaps 4;
QY 1 MGLSIFLLCVLGSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWLTA 60
Db 1 MLRLIALALVTGVHGET-RIIKGYECRPHSQPWQVAFQKTRLCCGATLIAPKWLTA 59
QY 61 AHCSGSRYVRLGHSLSQDWTQIRHSGFSVTHPGYLG--STSHEDLRLRLRPV 118
Db 60 AHCRKPHYVILLGHNLEKTDGCEQRWATSFPHDPFNNSLPNKHNRNDIMLVKMSVP 119
QY 119 RVTSVQPLPLPNDCATAGTCHVSGWGITNHPNPFPDLQLCLNLSIVSHATCHGVY 178
Db 120 PFTRAVQPLTLSPHCVAAGTSLISGWGTTSPQLRPLHSLRCANVSIIEHKECEKAY 179
QY 179 RITSNMVCYAG-GVPQDQACQDGGGGLVCGGVLQGLVSGVSGVPGQDGIPIGVYTYICKY 237
Db 180 NITDMLCASVKEGKDCQDGGGGLVCGVNSQIQIIISWGQ-DPCAVTRKFGVYTKVCKY 238
QY 238 VDWIRMNRN 248
Db 239 FNWIHEVMNRN 249
RESULT 6
QYQYN3
ID QYQYN3 PRELIMINARY; PRT; 276 AA.
```

```
AC DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DE DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DE Hipostasin prostate type (Adult male tongue cDNA, RIKEN full-length
DE DE enriched library, clone:2310015108 product:protease, serine, 20, full
DE DE insert sequence) (Adult male tongue cDNA, RIKEN full-length enriched
DE DE library, clone:2310040F07 product:protease, serine, 20, full insert
DE DE sequence).
DE GN PRSS20.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Prostate;
RC Yamauchi N., Mitsui S.;
RA Yamauchi N., Mitsui S.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Prostate;
RC Yamauchi N., Mitsui S.;
RA Yamauchi N., Mitsui S.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Tongue;
RC Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa M., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Tongue;
RC The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RN Nature 420:563-573 (2002).
RN [5]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Tongue;
RC RIKEN FANTOM Consortium;
RA "Functional annotation of a full-length mouse cDNA collection.";
RN Nature 409:685-690 (2001).
RN [6]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Tongue;
RC MEDLINE=20499374; PubMed=11042159;
RX Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RN Genome Res. 10:1617-1630 (2000).
RN [8]

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Katsunai T., Tashiro H., Itoh M.,
RA Konno H., Akiyama J., Nishi K., Kitsumi M., Nishine T., Harada A.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Ikegami T., Kashiwagi K.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa K., Tanaka T., Tanaka T., Matsuda M.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000)
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AB016227; BAA36955.1; -;
DR EMBL; AK009360; BAB26241.2; -;
DR EMBL; AK009720; BAB26461.2; -;
DR HSSP; P00763; IDPO.
DR MEROPS; S01.257; -;
DR MGD; MGI:1929977; Prss20.
DR GO; GO:0005576; C:extracellular; IDA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 276 AA; 30753 MW; 90BDC03A8AB178D6 CRC64;

Query Match 43.5%; Score 597.5; DB 11; Length 276;
Best local Similarity 47.0%; Pred. No. 1.2e-49;
Matches 118; Conservative 33; Mismatches 95; Indels 5; Gaps 4;

Qy 1 MGLSIFLLLVGLSQATPKIFNGTECGNSOPWQVGLFEGTSLRCGVLIDHRWLTA 60
Db 28 MLRLIALALVTGVTGTTTNGYECRPHSQPQWQVQLKTLCCATLIAPKWLTA 86
Qy 61 AHCSGSRYWVRLGHEHSLQDWTETQIRHSGFSTHYPGLGA--STSEHDLRLRLP 118
Db 87 AHCRKPHVILLGEHLEKTDGCEQRMMATESPFPDFNNSLPNKDHNDIMLVKMSPV 146
Qy 119 RVTSVQPLPLPNDACATAGTECHVSGWGIHNPENPEFDLLQCLNLSIVSHATGVP 178
Db 147 FFTRAVQPLTLSPHCVAAGTSLISGWGTTSSPOLRPHSLRCANVSIIEKECEKAY 206
Qy 179 RITSNMVCAG-GVPGQDACQDSCGGLVCGVGLQGLVSGVSGVPCGQDGPVGYTYCK 237
Db 207 NIITDMLCASVRKEGKSCQDSCGGLVCGVGLQGLIISWGQ-DPCAIVTRKPGVTK 265
Qy 238 VDWTRMTMNRN 248
Db 266 FNWIEHVMNRN 276

RESULT 7
ID Q96RQ0 PRELIMINARY; PRT; 255 AA.
AC Q96RQ0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Prostinogen.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21226193; PubMed=11327827;
```

```
RA Takayama T.X., Carter C.A., Deng T.;
RT "Activation of prostate-specific antigen precursor (pro-PSA) by
RT prostin, a novel human prostatic serine protease identified by
RT degenerate PCR";
RL Biochemistry 40:1679-1687(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF303046; AAK62813.1; -.
DR HSSP; P00761; IAN1.
DR MEROPS; S01.081; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYSPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 255 AA; 27986 MW; 00D5B79E14B9468F CRC64;

Query Match 43.3%; Score 595; DB 4; Length 255;
Best Local Similarity 46.5%; Pred. No. 1.9e-49;
Matches 119; Conservative 32; Mismatches 87; Indels 18; Gaps 6;

QY 8 LLVLGLGQAAT---PKIFNCTECGRNSQPMQVGLFECTSLRCGVLIDHRWVLTAAHC 63
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3 LLLTSLFLLTAQDQKLEDEGAPHSQPMQVGLYGRNCGASLSPHWVLSAAHC 62
QY 64 SGRSYVWVRLGSHLSQLDWTQTEIRHSGFVTHPGYLGASTSHEHDLRLRLPLVRVTSS 123
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 QSRFMRVRLGHNLRKRDGPEQLRTTSRVIPHRY--EASRNDIMLLRLVQPARLNQ 120

QY 124 VQPLPLPNDCATAGTECHVSGMGIYNH-----PRN--PPDLQLCLNLSIVSHATCH 173
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 VRPAVLPTCRPHGACVVGSLVSHNEPFGTAGSPRSQVSLPDTLHCANIGIISDTS 180

QY 174 GYVPGRTSNMVCAGV-PGODACOGDSGGLVCGVGLGVLGWSVSGVPGQDGIPIGVYT 232
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 KYPRLNTMVCAGAEGRGAECSGDSGGLVCGVGLGVLGWSVSGVPGDVTTEGVYT 239

QY 233 YICKYVDIMRMN 248
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
240 KVCHYLEWIRETMKN 255

RESULT 8
Q9D140 PRELIMINARY; PRT; 293 AA.
ID Q9D140;
AC Q9D140;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 1110030019Rik protein.
GN 1110030019Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giasi C., King B., Kuchiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
```



```
Query Match 42.1%; Score 578; DB 11; Length 242;
Best Local Similarity 49.1%; Pred. No. 7.9e-48;
Matches 114; Conservative 35; Mismatches 75; Indels 8; Gaps 5;

QY 21 KIENGTEGRNSQPMQVCLFEGTSLR--CGGVLDHRLVWLTAAHCSGRYVRLGHSLS 78
D 15 KIIIGYRCVRNSQPMQVALQPGHRLFCGGVLLSDQWVITAAHCAIRLILHVALGKHNR 74
QY 79 QLDWTEQIRHSGFSVTHPGYLCASTSHSHEDLRLRLRLPVRVTSVQPLPLPNDCATAGT 138
D 75 RWEATQVVRVARQVPHQY--QPOAHNDLMLLKLQKVALGRAVKTVISVASSCASP 132
QY 139 ECHVSGWGITHNRPFPDQLQCLNLSIVSHATCHGVIPGRITSNMVCAGVGP--GDAC 196
D 133 PCRVSGWGTIASPIARYPTALQCVNVMIMSEQACHRAYPEGIIITSGMVCA-GVPEGGKDC 191
QY 197 QGDSGGLVCGGVQLQGLVSGVSGPCGQDGIPGVYTYICKYVDWTRMIMRNN 248
D 192 QGDSGGLVCGGVQLQGLVSGW-MERCAMPGPVGVYANLONVHSWIQRTMQSN 242

RESULT 12
Q7SZT1 PRELIMINARY; PRT; 248 AA.
AC Q7SZT1;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Trypsin (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2];
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner B.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3];
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Klein S., Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC056068; AAH56068.1; -.
```

```
FT NON TER 1
SQ SEQUENCE 248 AA; 26494 MW; 1ED9D0D218702860 CRC64;

Query Match 41.0%; Score 563; DB 13; Length 248;
Best Local Similarity 45.0%; Pred. No. 2.3e-46;
Matches 112; Conservative 46; Mismatches 83; Indels 8; Gaps 5;

QY 3 LSIFLLCVLGLSQA--ATPKIFNGTEGRNSQPMQVCLFEGTSLRCCGVLDHRLVWLT 60
D 5 MKELVILVLLGAAVAFEDDDDKIVGFTCAKNAPVQVSLNAGYHF-CGSLINSQWVVA 63
QY 61 AHCSSRYVWVLGHSLSQLDWTQIRHSGFSVTHPGYLCASTSHSHEDLRLRLRLPVRV 120
D 64 AHCYKSRITQVRLGHNIALNEGTEQFIDSKVIPHNY--NSRNLNDIMLIKLTARL 121
QY 121 TSSVQPLPLPNDCATAGTECHVSGWGITHNRPFPDQLQCLNLSIVSHATCHGVYPCR 180
D 122 SANTQSVPLPSACASAGTNCILSGWNTLSSGTWYPDLLQCLNAPILTDSQCSNYPGEI 181
QY 181 TSNMVCAGVGP-GDACAQDGGPLVCGGVQLQGLVSGVSGPCGQDGIPGVYTYICKYVD 239
D 182 TKNMFCAGFLAGKXDCQDGGPVCNGQLQGVVSGY--GCAQRNYPGVYTKVCFNEVT 239
QY 240 WIRMIMRNN 248
D 240 WIQSTISSN 248

RESULT 13
Q7SZ06 PRELIMINARY; PRT; 243 AA.
AC Q7SZ06;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2];
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
```

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RC SEQUENCE FROM N.A.
RC TISSUE=Whole;
RA Klein S., Strausberg R.; to the EMBL/GenBank/DBJ databases.
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC054194; AAH54194.1; -.
KW Hypothetical protein.
SQ SEQUENCE 243 AA; 25476 MW; C5B82B458B209F31 CRC64;

Query Match 40.4%; Score 555; DB 13; Length 243;
Best Local Similarity 45.7%; Pred. No. 1.3e-45;
Matches 113; Conservative 42; Mismatches 82; Indels 10; Gaps 6;

Qy 6 FLLLCVLGLSQAAT--PKIFNGTECGRNQPMQVGLFEGTSLRCGGVLIDHRLVLTAAH 62
Db 3 FLLLCVL-LGAAAFDDDKIIGATCAKSSVPYIVSLNSGVHF-CGSLITNQWVSAAH 60

Qy 63 CSGRYWVRLGEHSLDQTEQIRHSGFVTHPGYLGASTSHEHDLRLRLPVRVTS 122
Db 61 CYKASIQVRLGEHSLDQTEQIRHSGFVTHPGYLGASTSHEHDLRLRLPVRVTS 118

Qy 123 SVQPLPNDPCATAGTECHVSGWGLTNHPRNPFDPDLQCLNLSIVSHATCHGVYVPGRTS 182
Db 119 AVNTVPLPSGCAAGTSLISGWNLTLSNGNYNPDLLQCLNAPILTNAAQNSAYPGEITA 178

Qy 183 NMVCGAGVP-QGDACQSDSGPLVCGGVLOGLVSGSVGPGCGQDIPGVYTYICKYVDWI 241
Db 179 NMICVGFMEGKDSQCGSGGPPVCGNLQGVWSGY--GCAMRNYPGVYTKVCYNNAWI 236

Qy 242 RMIMRN 248
Db 237 QNTIAAN 243

RESULT 14
O88301 PRELIMINARY; PRT; 246 AA.
ID O88301
AC O88301;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serine protease (BSP).
GN PRS18 OR MBSP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Matsui H., Takahashi T.;
RT "A novel serine protease preferentially expressed in brain.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C; TISSUE=Skin;
RA Meier N., Dear T.N., Boehm T.;
RT "A novel serine protease expressed in the hair follicle.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL;
RA Matsui H., Takahashi T.;
RT "The sequences of mouse serine protease gene expressed in brain.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AB015206; BAA28895.1; -.
DR EMBL; Y18723; CAA7269.1; -.
DR EMBL; AB032402; BAA84544.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.236; -.
DR MGD; MGI:1343166; Prss18.
DR GO; GO:0008233; F:peptidase activity; IEA.

GO; GO:0004263; F:chymotrypsin activity; IEA.
GO; GO:0008233; F:peptidase activity; IEA.
GO; GO:0004295; F:trypsin activity; IEA.
GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
InterPro; IPR009003; Cys Ser trypsin.
InterPro; IPR001254; Peptidase S1.
InterPro; IPR001314; Peptidase_S1A.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
SMART; SM00020; Tryp_SPC; 1.
PROSITE; PS00240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 246 AA; 27496 MW; E20C080087139B63 CRC64;

Query Match 40.3%; Score 553.5; DB 11; Length 246;
Best Local Similarity 43.5%; Pred. No. 1.9e-45;
Matches 107; Conservative 46; Mismatches 82; Indels 11; Gaps 6;

Qy 7 LLLCVLGLSQA-----TPKIFNGTECGRNQPMQVGLFEGTSLRCGGVLIDHRLVLTAAH 62
Db 4 LALCLV-LAKSAWSEEQEKVHGCPCLKDSHPFQAALYTSCHLLCGGLIDPQWVLTAAH 62

Qy 63 CSGRYWVRLGEHSLDQTEQIRHSGFVTHPGYLGASTSHEHDLRLRLPVRVTS 122
Db 63 CKKPNLVQLGKHLRQTETFRQISVDRITVHPRY--NPETHNDIMVHLKNPVKFSK 120

Qy 123 SVQPLPNDPCATAGTECHVSGWGLTNHPRNPFDPDLQCLNLSIVSHATCHGVYVPGRTS 182
Db 121 KIQLPLKNDCEBENPCQLLWG--KXNGDFDPTIQCADVHLVPREQCERAYPGKITQ 178

Qy 183 NMVCGAGVP-QGDACQSDSGPLVCGGVLOGLVSGSVGPGCGQDIPGVYTYICKYVDWI 241
Db 179 SMVCGADMKEGNSDQCGSGPLVCGGRLRLGLVSWGDM-PCGSKEKPGVYTDVCTHRWI 237

Qy 242 RMIMRN 247
Db 238 QNILRN 243

RESULT 15
Q91Y82 PRELIMINARY; PRT; 253 AA.
ID Q91Y82
AC Q91Y82;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Neurosin (Protease, serine, 18).
GN PRS18.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Mitsui S., Yamaguchi N.;
RT "cDNA cloning and characterization of mouse brain specific serine
protease, Neurosin.";
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AB008928; BAB5605.1; -.
DR EMBL; BC031119; AAH31119.1; -.
DR HSP; P00761; 1ANI.
DR MEROPS; S01.236; -.
DR MGD; MGI:1343166; Prss18.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.

```



